RGBD-based Robot Localization in Sewer Networks

D. Alejo*, F. Caballero† and L. Merino*

***Abstract*— This paper presents a vision-based localization system for global pose estimation of a sewer inspection robot given prior information of the sewer network from local institutions. The system is based on a Monte-Carlo Localization system that uses RGBD odometry for the prediction stage. The update step takes into account the sewer network topology for discarding wrong hypotheses. Moreover, this step is further refined whenever a discrete element of the network (i.e. manhole) is detected. To this end, another RGBD camera pointing upwards is used for precise manhole detection. A Convolutional Neural Network has been successfully trained for classifying images with and without manholes with 96% accuracy over the tested dataset. The complete system has been validated with real data obtained from the sewers of Barcelona yielding accurate localization results. All the logs and code used in the context of this paper are publicly available.**

I. INTRODUCTION

Sewer inspections require many people to work in risky and unhealthy conditions. Sewers are classified as confined spaces which require special health and safety measures, and pose additional risks like slippery floors, obstacles and biological hazards from the potential contact with wastewater. These features make the process of sewer inspection a risky and expensive process that requires improvements. Therefore, introducing a robotic solution in this process aims at reducing the labor risks, improving the precision of sewer inspections and optimizing sewer cleaning resources of the city.

In particular, the work described in this paper has been carried out within the framework of the Challenge "Utility infrastructures and condition monitoring for sewer networks. Robots for the inspection and the clearance of the sewer network in cities [1] inside the Echord++ EU project. This challenge proposes the use of autonomous robots for the inspection of sewers in Barcelona. The sewer network of Barcelona is $1,532$ km. long, from which approximately 50% is accessible, which means that the pipe is at least 1.5 m. high and workers are allowed to go inside. The requirements of the challenge is to design and demonstrate a semi-autonomous robot which has to be operated from surface. Also, the robot should provide precise positioning of structural defects and other alerts in global coordinates. Therefore the development of a localization system that allows the operator to keep track of the robot position is a key feature in the development of the project.

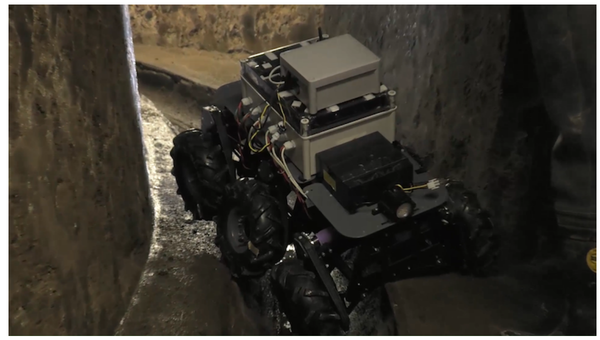

Fig. 1. The prototype used in the field experiments presented in this paper.

Fig. 2. Two snapshots taken from the frontal camera of the robot at different localizations.

Ground robots are the current preferred solution for this task, given the requirements on robot autonomy, payload and the environment. There is a wide variety of robotic platforms for gallery inspection already in the market. A survey can be found in [1]. However, most of them have been designed in order to inspect small pipes and sewers. Therefore, they cannot be used within the project as they cannot cover the minimum inspection distance requirement of more than 1 km per labor day. Moreover, they are connected to the base station by means of a wire, while the project requires a wireless solution. To fulfill these requirements, we propose the use of a new platform named Sewer Inspection Autonomous Robot (SIAR). It is a six-wheeled ground robot specially designed for sewer inspection (see Fig. 1). The current prototype has already been deployed in real sewers in several experiments.

We use RGBD cameras as the main sensor modality. They offer some advantages in this particular environment, as they provide directly range information even in dark conditions, do not need mechanical moving parts and are cheap sensors. The robot carries a sensor head consisting of 5 RGBD cameras. Two images captured by the frontal camera are shown in Fig. 2.

In this paper, we present a system that combines visual

This work was partially supported by the ECHORD++ project with grant agreement no. 601116

*D. Alejo and L. Merino are with School of Engineering, Universidad Pablo de Olavide, Seville, Spain. Email: {david, lmercab} at upo.es

†F. Caballero is with the Department of Systems Engineering and Automation, University of Seville, Spain. Email: fcaballero at us.es

[1]http://echord.eu/pdti/pdti-urban-robotics-sewer-inspection/

odometry (VO) with the recognition of manholes in depth images for global sewer localization. This recognition is based on machine learning. The use of the manhole as main features for localization is done for three main reasons. First, there exists a regulation of the maximum distance between two consecutive manholes for safety reasons. Second, their diameter (70 cm) enables their detection with an acceptable success rate. Third, their positions are labeled in the Global Information System (GIS) provided by the local agencies.

The paper is organized as follows. Next section summarizes related work on sewer localization. Section III overviews the main components of the system. Section IV describes the machine learning method for manhole detection, and Section V describes the localization method. The paper concludes with experimental results and conclusions.

II. Related work

The localization of a robot in a sewer system has a number of issues to be considered. Most systems rely at first stage in odometry information acquired from wheel encoders and inertial units. However, humidity, water and waste significantly decrease the wheel grip, distorting the computed linear and angular velocities based on wheel encoders. Besides, the robot has often to negotiate steps in order to cross channels, which can further distort the measures obtained from the wheel encoders. Even though some devices can improve the odometry reliability [2], the errors cannot be ignored in the long term.

One of the first localization systems, for the KURT inspection robot, presented in [3], already recognizes the main problems, like wheel slippage for odometry. Topological localization is done by recognizing junctions from data and matching them to a map. Wheel odometry is considered in [4], where also junctions are classified using readings from wheels and used to compensate the errors. Only simulations are considered.

Alternative methods to wheel odometry have been proposed. In [5], the fading period of RF signals in tunnels (assimilated as waveguides) are used to obtain an odometry-like measurement for localization. LIDAR-based systems are also considered for motion estimation [6]. An additional problem in this case is the lack of features and the symmetry of the environment, which makes more difficult to estimate certain degrees of freedom. Tethered robots can use also the tether to estimate the motion of the robot [7]. The same paper considers also VO for robot motion estimation. The visual odometer is in this case based on the known geometry of a cylindrical sewer. Here we consider a model-free VO approach, as the sections of the sewer are not uniform.

Besides odometry estimation, many inspection applications in sewers require global localization into the sewer. Some systems are based on the communication of extra-low frequency signal with an external mobile system [8] to estimate the position of the robot. However, this requires additional infrastructure, and cannot be applied to all types of sewers.

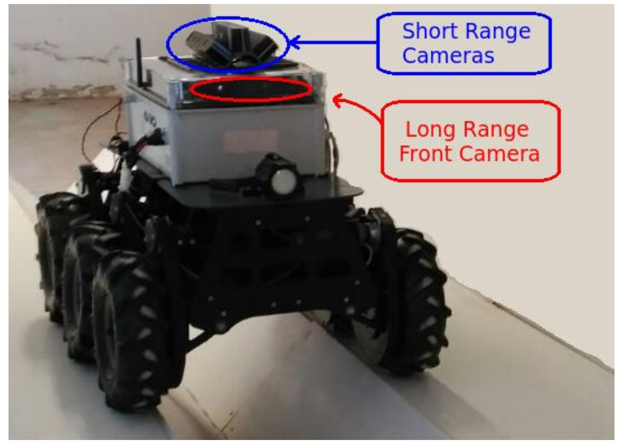

Fig. 3. Proposed layout of the cameras on top of the platform. Note that the frontal camera is inside the cover

A localization system is also presented in [7], based on matching the same visual features used for VO stored in a map. However, results are obtained in a small sewer testbed. Many times there are previously available maps from Geographical Information Systems, containing the positions of elements of the sewer network, like manholes, inlets and other elements. Here we consider an hybrid metric-topological Monte-Carlo localization method in which VO errors are compensated by recognizing those elements.

III. Proposed system

In this section we briefly summarize the robotic platform employed. Then we present the sensors used for localization, as well as a general overview of the approach.

A. Robotic platform

Figure 1 presents the current robotic platform used in real experimental scenarios. The robotic frame is made with IP67 enclosings in order to accommodate for the hardest environmental conditions during sewer inspection. A six-wheeled ground robot with independent traction has been designed and implemented by the company IdMind[2] as part of this work. It is able to navigate over a wide range of floors and small obstacles, including steps of over 20 cm. The axis of the two central wheels is positioned below the other ones, and the center of gravity of the robot is located over that axis, so that for turning the robot can behave closely to a differential-drive robot. This configuration is very flexible and adapts very well to the different scenarios the robot can find during routine inspections. It carries LiFePo batteries for batteries and electronics, with an autonomy of 5 hours. For more details on the platform, please refer to [9].

B. Sensor system

The platform is equipped with 5 RGBD cameras (see Fig. 3). Two long-range RGBD cameras (with a maximum range of 8 m) are located in the front and rear sides of the robot. These cameras are used for localization, navigation,

[2] `http://www.idmind.pt`

obstacle avoidance and 3D reconstruction of the ground and the walls. The remaining three sensors have been placed in order to enhance the detection of possible structural defects at the ceiling and in the upper parts of the walls and for the detection of manholes. The system is equipped with front and rear lights that are vital for providing the operator with a clear view of the sewers and for using VO.

C. Localization System

The proposed localization system is based on the integration of VO, detection of manholes and Monte-Carlo localization. A brief description of the main functional blocks of the proposed approach is introduced below.

- **RGBD Odometry**. While the robot provides wheel odometry, due to the typical slippage mentioned above, the values are not reliable. Thus, the main odometry source is a VO method [10] developed for stereo-vision and adapted to RGBD. This block takes as input the RGB and depth flows of the frontal camera of the robot, matches robust features from consecutive or closely spaced frames and then obtains the relative pose between cameras that minimizes the re-projection error.
- **Manhole detector**. This module will check whether the robot is under a manhole or not based on the depth images gathered by a camera pointed upwards (see Fig. 3). A machine learning approach is employed to perform this classification robustly and computationally efficient. It is described in Section IV.
- **Localization module**. This module integrates the odometry measurements thought time, performing proper corrections according to an *a priori* topological map and the manhole detector. In this paper, a Monte-Carlo localization system [11] has been used (see Section V).

IV. Machine learning for automatic manhole detection

The high symmetry of the sewer gallery makes localization a very complex problem. Commonly used tools for localization as visual place recognition algorithms usually suffer from low accuracy when targeting repetitive visual structures [12]. This problem is accentuated by the poor illumination conditions and high symmetry of the environment. On the other hand, detailed 3D/2D maps [13][14] of the sewer network are frequently unavailable due to its vast size.

Manholes offer a good opportunity for localization. They use to be well localized in the sewage system, and also in the city drawings. This information can be used to reset the localization errors if the robot is able to identify the manhole on top of it. In addition, manholes also have a particular shape, they break with the uniformity of the gallery ceiling, which might simplify its detection (see Fig. 4).

Thus, our objective is to automatically detect if the robot is lying under a manhole. We use the depth images provided by a camera placed on top of the robot, pointing toward the ceiling. Our approach makes use of machine learning approaches to train a classifier able to split the depth camera view between regular ceiling and manhole.

Fig. 4. View of the sewer gallery from the depth camera pointing towards the ceiling. (Left) Regular gallery ceiling. (Right) Manhole

Considering the significance of these detections for the localization system, we impose the following two hard constrains: the accuracy must be very high to avoid false positive or false negative detections that might impact in the localization results, and second, the manhole detection must be light-weight from the computational point of view to not compromise the rest of the navigation task.

With all the previous constraints in mind, Convolutional Neural Networks (CNNs) emerge as a convenient technique as they have been extensively used for image classification purposes in last years with great success. Their ability to extract hight level features from the images and to easily reduce the dimensionality of the problem are key factors for their success. In addition, this dimensionality reduction helps to decrease the computational requirements of the classifier.

Next paragraphs summarize the designed CNN architecture, the training process and the validation.

A. Convolutional neural network architecture

The CNN architecture developed in this paper is similar to the AlexNet architecture [15]. In these networks we combine convolution layers with pooling layers and a fully-connected network. One of the advantages of CNNs over Multi-Layer Perceptrons (MLP) is that convolution layers are able to extract constant features from inputs. A convolutional layer is a compound of kernels or filters with a certain dimension (mostly depending on the input resolution). These filters convolve with different receptive fields on the input, going all over the image. Thus, each convolution layer unit receives the information of a small region of the previous layer with the kernel size. The result of a convolution operation is a feature map that can detect a particular spatial aspect (edges, corners, etc.). By combining several convolution layers the network is able to extract higher-order features.

The CNN architecture designed for manhole detection is shown in Fig. 5. The first step consist in down-sampling the depth image to a smaller scale, 80×60 particularly. We perform this down-sampling because we assume most of the details of the manhole are retained in the reduced version, while it decreases the computation. The network is composed by four convolutional layers followed by ReLU activations and MaxPooling, except for the last one. We can see how we reduce the size of the inputs until having a tractable amount of information. The output of the fourth convolution layer is fed into a fully-connected layer of 100 neurons and ReLU activations. Finally, the output layer is composed by a single fully-connected neuron with sigmoid

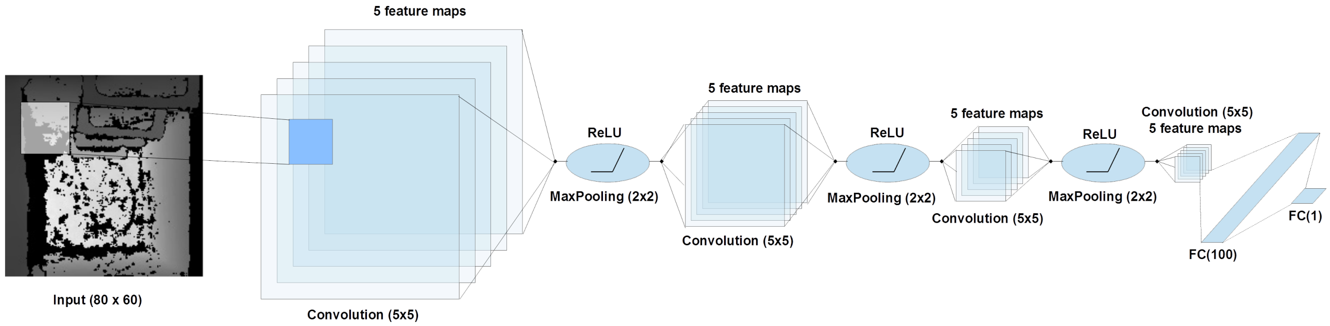

Fig. 5. Convolutional Neural Network for automatic manhole detection based on depth images

activation functions. The result is a binary value that is true whenever a manhole is detected and false otherwise.

Notice in Fig. 5 how the number of filters (feature maps) for each convolutional layer is just five. This small number of filters is taken on purpose, in order to keep the number of parameters under control. Thus, the total number of parameters of this CNN structure is 8,921, of which 7,600 belongs to the first fully connected layer. A CNN prediction is performed in less than 1.5 ms in a regular i7 with two Cores thanks to this small number of parameters.

B. *Dataset for training*

We need correctly labeled depth images gathered by the robot to perform the training of the proposed CNN. To this end, an experimental campaign was developed in real sewers in the city of Barcelona. The data set was obtained from two experiments (see Section VI). We used Experiment 1 for training and Experiment 2 for validation.

In both datasets, training and validation, a manual labeling of the depth images with manholes was done. The positive set of the training dataset was expanded by introducing horizontal and vertical flips, and small translations and rotations.

As a whole, the dataset is composed by $40,000$ depth image samples with resolution 80×60. Each sample integrates a label that indicates if the image contains a manhole or not. From these samples, $21,000$ are used to to validate the CNN and are not included into training process.

C. *Training*

We implemented the networks using Python and the Keras library [16]. The learning algorithm used to train the networks is the Adaptive Moment Estimation (Adam), an algorithm for first-order gradient-based optimization. It has been demonstrated the effectiveness of Adam over other stochastic first-order methods on MLP and CNN [17].

Each learning process takes 20 epochs with a batch size of 100 samples. A 20% dropout layer [18] was included after every convolutional or fully connected layer (except for the output) to prevent the computed neural networks from overfitting.

D. *Validation*

After the training process, we validate the model with a set of $21,000$ samples not used in the training (test set). The obtained accuracy results show a 99% of accuracy in the training dataset whereas roughly a 96% in the validation one. These results are very satisfactory, taking into account that different types of galleries visited in both datasets and that there were opened and closed manholes.

Table I summarizes the true and false positive and negative rates of the network. Note that false positive rate is approximately of a 2%. On the other hand, the false negative rate is bigger, going up to a 10%. This indicates the probability of missing one manhole by considering just one image. However, the robot can take tens of images of the manhole as it traverses sewer. Therefore, the probability of missing a manhole can be reduced during robot execution by applying basic temporal filtering of the detection outcome.

TABLE I

MANHOLE DETECTOR CONFUSION MATRIX

		Predicted	
		Positive	Negative
Actual	Positive	0.90	0.10
	Negative	0.02	0.98

V. GRAPH-BASED LOCALIZATION

As previously introduced, the approach is based on Monte-Carlo Localization [11], which makes use of a particle filter to represent the robot localization belief.

In the proposed filter, each particle represents a hypothesis which consists of a 2D position with orientation. The z coordinate can be obtained from the topological map as in [19] but this is beyond the scope of the paper. The hypotheses are validated (weighted) according to the position of each particle compared with a topological map obtained from GIS data. Figure 6 shows the GIS data used in the experiments to build the graph for the localization module. This graph contains manhole vertices, where a manhole can be detected; and fork vertices, in which several sewers converge, or a sewer turns into another direction. These vertices are linked by edges that indicate traversable paths between them. The next subsections detail each step of the filter.

A. *Initialization*

This step is currently done manually at the location of the manhole where the robot has been deployed. The output

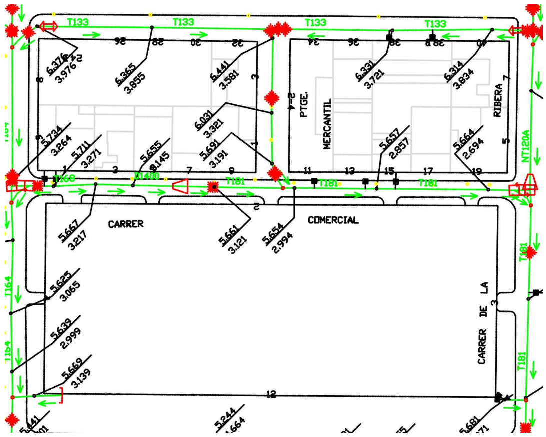

Fig. 6. Topological map used as input for generating the internal graph in the localization module

of this step is the initial population of particles, which are located randomly with a multivariate normal distribution centered at the user defined location.

B. Prediction

The position hypotheses (particles) are predicted based on the visual odometry from the RGBD cameras. Note that the odometry measures are modified with independent additive Gaussian noise and thus each particle will evolve differently.

C. Update

The filter also needs a method to validate that particles are in the right position. In this paper, we propose a two step validation process, that will weight the particles differently whether a manhole is detected or not.

1) Edge weighting: measuring lateral errors: When no manhole has been detected in the current image, the particles are ranked according to their distance to the closest edge of the topological map. Particularly, we use the criteria:

$$w_{edge} = \frac{1}{\sigma_e\sqrt{2\pi}} e^{-\frac{d_e^2}{\sigma_e^2}} \tag{1}$$

where d_e is the distance from the particle to the closest edge and σ_e is related to the width of the sewer.

It is worth to mention that the map could not be as accurate as desired and thus it could be a good idea to overestimate the edge deviation. This is more evident in the surroundings of forks and turns. Therefore, in this paper two different values of the standard deviation have been used depending on the proximity of a non-straight section.

2) Manhole weighting: measuring longitudinal uncertainty: Navigating through the sewer environment will always increase the longitudinal uncertainty in the position of the robot. In this paper, manholes are detected to reduce this uncertainty and precisely localize the robot. Eq. 2 is used to rank the particles of the filter whenever a manhole is detected by the detection module.

$$w_{manhole} = \frac{1}{\sigma_m\sqrt{2\pi}} e^{-\frac{d_m^2}{\sigma_m^2}} + w_{manhole}^{min} \tag{2}$$

where d_m is the distance from the particle to the closest manhole in the graph and σ_m is the considered standard deviation surrounding the manhole, and is related to the size of the area were they can be detected. The additional term $w_{manhole}^{min}$ term has been added to the weight in order to mitigate the effects of a false detection in places far away of manholes. In such cases, the following in-equation holds for all particles:

$$w_{manhole}^{min} >> \frac{1}{\sigma_m\sqrt{2\pi}} e^{-\frac{d_m^2}{\sigma_m^2}} \tag{3}$$

Therefore, all particles are uniformly weighted in this case. On the other hand, in the surroundings of the manhole the most important term should be due to the normal distribution:

$$w_{manhole}^{min} << \frac{1}{\sigma_m\sqrt{2\pi}} e^{-\frac{d_m^2}{\sigma_e^2}} \tag{4}$$

This equation must hold if $d_m < d_{detect}$, where d_{detect} is the maximum detection distance of a manhole.

D. Resampling

Whenever the dispersion of the particle set exceeds a threshold, or the number of maximum number of updates is reached, the set of particles is restructured through importance resampling. The new set of particles is obtained by randomly sampling the old set with a sampling distribution proportional to the weights of the particles. We make use of the low variance sampler described in [20]. By periodically resampling the particles according to the weights proposed in the previous sections, we ensure that most of the particles are located according to the prior GIS information.

However, false positives in the surroundings of a manhole can make the particles of the filter to be resampled into an area distant from the real position of the robot. To mitigate such effects, we only perform manhole updates when more than M_{thres} manhole measures are received between two consecutive update intervals.

VI. Field Experiments

This section discusses the results of the proposed system when applied to field data. Firstly, the experimental scenario is described. Readers are encouraged to reproduce the results, as the code and data of the localization system can be openly accessed on [3]. There, instructions for running the two experiments and to obtain the dataset used in this paper can be found. All the code is developed under the Indigo distribution of the Robotic Operating System (ROS) [4].

A. Scenario

The sewers considered for the experiment are located in the surroundings of the Mercat del Born, Barcelona (see Fig. 7). During the experiments, the robot was teleoperated from a short distance in Experiment 1 and Experiment 2.

[3]https://github.com/robotics-upo/siar_packages
[4]www.ros.org

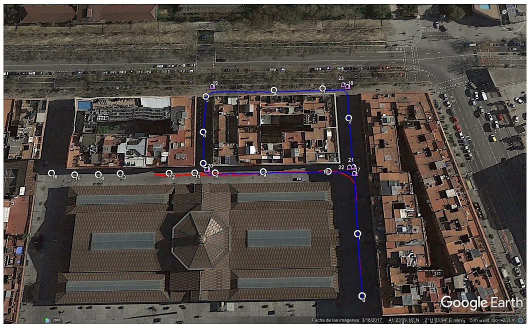

Fig. 7. Scenario where the experiments have been carried out. The depicted lines represent an example of localization results in Experiments 1 (blue) & 2 (red). Manholes are labeled in a white circle and forks in purple.

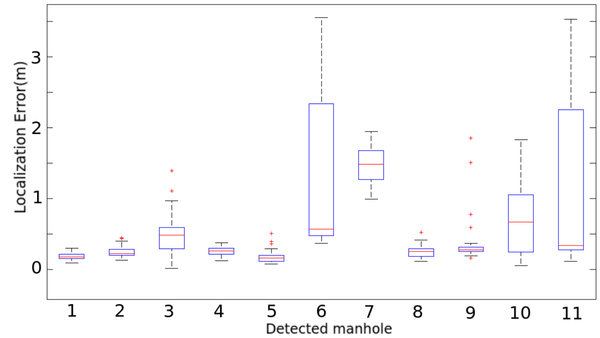

Fig. 8. Distribution of the distance from the mean of the particle set to the ground truth position of the manhole in Experiment 1.

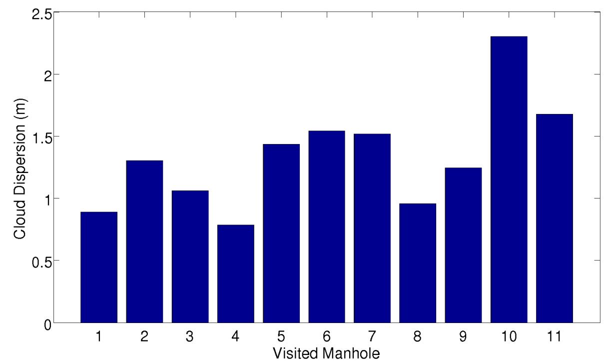

Fig. 9. Mean dispersion of the particles in the instants where the robot passes just under a manhole in Experiment 1.

B. Localization results

The complete localization system has been successfully tested with the data gathered in the sewers. As the algorithm is randomized, the results presented here are calculated taking into account the output from 30 different executions, all of them using the same input data from both Experiment 1 and 2. The traveled distance in these experiments was of about 400 m.

Figure 7 shows an example of the trajectories of the robot in Experiments 1 & 2 as estimated in an execution of the proposed localization algorithm. Please note that the obtained trajectories are smooth and strictly follow the guidelines imposed by the topological map that is loaded as input.

One of the main difficulties when analyzing the behavior of a localization system is the (in)existence of ground truth data for making comparisons with the obtained results. In this paper, no ground truth is available for the robot trajectory. Thus, basic statistics of the particle set (namely mean position and standard deviation in the three coordinates) when the robot is just under a manhole are saved. These instants have been manually labeled according to the images of the camera pointing upwards. Then, the results of the mean position are compared to the GPS position of the manhole. The localization is considered accurate when the distance between the mean and the ground truth position is below the distance for manhole detection (50 cm. approx.).

1) Experiment 1: Figure 8 represents the distribution of the distances of the weighted mean of the set of particles to the ground truth position of the manhole above the robot in each execution of the algorithm. In this experiment, the robot follows a track with a loop and passes through 11 manholes, revisiting only one of them (the first and tenth ones). The main result is that the localization algorithm succeeded to accurately localize the robot in all of the executions. Moreover, results show that the mean distances of the obtained particle sets to the actual position of the manhole rarely exceed 1 m. whenever the robot was under a manhole. In the cases where this deviation is high, it is probable that further manhole detections were produced afterwards, reducing the distance to the actual position of the robot.

In addition, it is important to analyze the evolution of the dispersion of the particle clouds as a the experiment goes on. Figure 9 represents the mean standard deviation of the particles when the robot is below a given manhole. As expected, the dispersion is closely related to the mean distance to the manhole, as larger localization errors are usually due to cumulative odometry errors that are reflected in the dispersion of the particles. It is also noticeable that taking into account the dispersion of the cloud and the mean distance to the manholes, a manhole resampling procedure performed would always find particles under the correct manhole and thus the proposed localization system would successfully reduce the uncertainties in the pose the robot.

2) Experiment 2: This experiment is slightly shorter than the Experiment 1. However we have found it to be more challenging due to the following reasons:

1) Only two 90 degrees turn carried out at the same places, while Experiment 1 included up to five turns of this kind. This helped to re-localize the robot, as the longitudinal uncertainty is reduced due to the eventual elimination of particles turning in wrong places.
2) A manual 180 degrees turn is carried out to make the robot go back through the sewer. This is a challenge for the VO system.
3) Navigation experiments were conducted after visiting manhole 3. The robot was to be recovered manually in some of them.
4) The speed of the robot was increased, making manhole

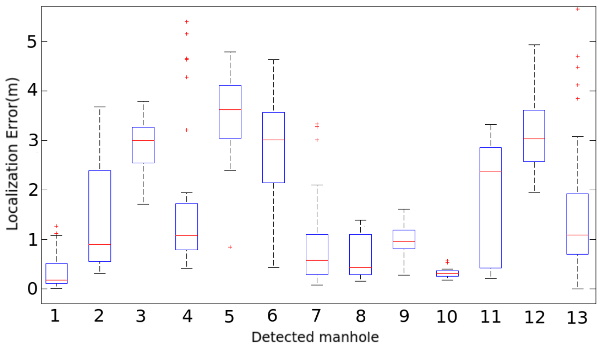

Fig. 10. Distribution of the distance from the mean of the particle set to the ground truth position of the manhole in Experiment 2.

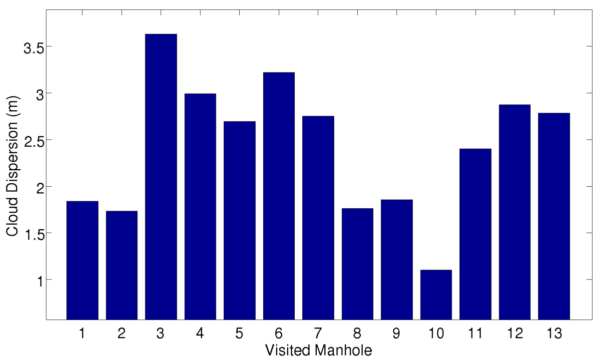

Fig. 11. Mean dispersion of the particles in the instants where the robot passes just under a manhole in Experiment 2.

detection more difficult.

Figure 10 represents the distribution of the distances of the weighted center of the set of particles to the ground truth position in Experiment 2. Due to the aforementioned causes, the distribution of the distances is worse than in Experiment 1. We expect that the introduction of further measurable elements can improve the results in this case.

The dispersion results presented in Fig. 11 are generally higher than the results of Experiment 1. This indicates that the localization obtained in this case is less accurate than in the previous experiment. This may also be produced because the experiment was also used for navigation tests in difficult areas, where the robot fell and had to be manually recovered.

VII. Conclusions and Future Work

A vision-based method for the localization of a ground robot navigating through real sewers has been designed, implemented and validated with real data. RGBD data is used for VO. Odometric drift is corrected using a map-based Monte Carlo localization system, using the known locations of manholes and sewer galleries as landmarks. A deep neural network has been trained and used to detect manholes using depth data from the RGBD cameras.

The paper describes results using data from real sewers in two different navigation experiments of 400 m. each. The experimental results indicate that the localization has been achieved in all executions even in the presence of sudden changes in the direction of the robot. Moreover, the median distance to the retrieved ground truth data is below 4 m. at the end of the experiments for all the executions.

As future work, the detection system can be easily trained to detect further elements of the sewer network, like inlets and others, which can be used to further refine the localization of the robot, and/or for inspection purposes.

References

[1] J. M. M. Tur and W. Garthwaite, "Robotic devices for water main in-pipe inspection: A survey," *Journal of Field Robotics*, pp. 491–508, 2010.

[2] J. N. Bakambu, V. Polotski, and P. Cohen, "Heading-aided odometry and range-data integration for positioning of autonomous mining vehicles," in *Proceedings of the 2000. IEEE International Conference on Control Applications. Conference Proceedings (Cat. No.00CH37162)*, 2000, pp. 279–284.

[3] J. Hertzberg and F. Kirchner, "Landmark-based autonomous navigation in sewerage pipes," in *Advanced Mobile Robot, 1996., Proceedings of the First Euromicro Workshop on*, Oct 1996, pp. 68–73.

[4] A. A. F. Nassiraei, M. Honda, and K. Ishii, "New approach to the self-localization of autonomous sewer inspection robots," in *2010 World Automation Congress*, Sept 2010, pp. 1–7.

[5] C. Rizzo, V. Kumar, F. Lera, and J. L. Villarroel, "'RF odometry for localization in pipes based on periodic signal fadings'," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sept 2014, pp. 4577–4583.

[6] D. Tardioli and J. L. Villarroel, "Odometry-less localization in tunnel-like environments," in *2014 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, May 2014, pp. 65–72.

[7] C. Walter, J. Saenz, N. Elkmann, H. Althoff, S. Kutzner, and T. Stuerze, "Design considerations of robotic system for cleaning and inspection of large-diameter sewers," *Journal of Field Robotics*, vol. 29, no. 1, pp. 186–214, 2012.

[8] H. Qi, X. Zhang, H. Chen, and J. Ye, "Tracing and localization system for pipeline robot," *Mechatronics*, vol. 19, no. 1, pp. 76 – 84, 2009.

[9] D. Alejo, C. Marques, F. Caballero, P. Alvito, and L. Merino, "SIAR: an autonomous ground robot for sewer inspection," in *Proceedings of the Spanish Actas de las Jornadas de Automática*, 2016, pp. 1–8.

[10] F. J. Prez-Grau, F. R. Fabresse, F. Caballero, A. Viguria, and A. Ollero, "Long-term aerial robot localization based on visual odometry and radio-based ranging," in *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, June 2016, pp. 608–614.

[11] S. Thrun, D. Fox, W. Burgard, and F. Dellaert, "Robust Monte Carlo localization for mobile robots," *Artificial intelligence*, vol. 128, no. 1-2, pp. 99–141, 2001.

[12] A. Torii, J. Sivic, M. Okutomi, and T. Pajdla, "Visual place recognition with repetitive structures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 11, pp. 2346–2359, Nov 2015.

[13] S. Oishi, Y. Jeong, R. Kurazume, Y. Iwashita, and T. Hasegawa, "ND voxel localization using large-scale 3D environmental map and RGB-D camera," in *2013 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Dec 2013, pp. 538–545.

[14] J. Biswas and M. Veloso, "Depth camera based indoor mobile robot localization and navigation," in *2012 IEEE International Conference on Robotics and Automation*, May 2012, pp. 1697–1702.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, p. 2012.

[16] F. Chollet, "Keras," https://github.com/fchollet/keras, 2015.

[17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014.

[18] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, Jan. 2014.

[19] J. Pérez, F. Caballero, and L. Merino, "Enhanced Monte Carlo Localization with Visual Place Recognition for Robust Robot Localization," *Journal of Intelligent & Robotic Systems*, vol. 80, no. 3-4, pp. 641–656, 2015.

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