Proceedings of 15th Int. Conf. on Intelligent Autonomous Systems,
Baden-Baden, Germany, June 2018

An Integrated Planning of Exploration, Coverage, and Object Localization for an Efficient Indoor Semantic Mapping

Diar Fahruddin Sasongko and Jun Miura

Department of Computer Science and Engineering,
Toyohashi University of Technology

Abstract. This paper describes an integrated viewpoint planner for indoor semantic mapping. Mapping of an unknown environment can be viewed as an integration of various activities: exploration, (2D or 3D) geometrical mapping, and object detection and localization. An efficient mapping entails selecting good viewpoints. Since a good viewpoint for one activity and that for another could be shared or conflicting, it is desirable to deal with all such activities at once, in an integrated manner. We use a frontier-based exploration, an area coverage approach for geometrical mapping, and object recognition model-based verification for generative respective viewpoints, and get the best next viewpoint by solving a travelling salesman problem. We carry out experiments using a realistic 3D robotic simulator to show the effectiveness of the proposed integrated viewpoint planning method.

Keywords: Viewpoint planning, semantic mapping, mobile robot.

1 Introduction

Mapping is one of the fundamental functions of mobile robots. A map is used for various purposes such as localizing a robot and finding a specific object. The process of mapping in an unknown environment can be divided into two aspects. One is *where to get* new infomration and the other is *how to integrate* all pieces of information into a consistent representation. The second one is related to so-called SLAM (simultaneous localization and mapping) methods and numerous approaches have been proposed [1] and matured in a certain range of mapping problems. The first one is related to *viewpoint planning* and is still a hot research area [2–4]. This paper focuses the first aspect in semantic map generation.

Viewpoint planning methods differ in objectives, that is, what information will be obtained from the observations at selected viewpoints. For example, pure exploration tries to increase the area of known space [2], while in object search, a robot selects a viewpoint to increase the probability of a detected object being a target object [5]. Multiple, sometimes conflicting, objectives can also be considered in an integrated manner [3].

Our previous viewpoint planner for geometric mapping [6] considers exploration and geometric data acquisition in an integrated manner. This paper extends the planner for semantic mapping, by additionally considering viewpoints for object verification. We show that an integrated viewpoint planning is better than non-integrated sequential planning by experiments in a realistic simulator.

The rest of the paper is organized as follows. Section 2 describes related work. Section 3 describes the target task and the overview of the proposed method. Section 4 briefly explains our previous planner. Section 5 describes object recognition and viewpoint genereation for object verification. Section 6 shows experimental results to compare the sequential and the integrated approaches. Section 7 summarizes the paper and discusses future work.

2 Related Work

2.1 Exploration

Exploration of an unknown environment has been an important ability of mobile robots. In usual mapping algorithms, a whole environment is divided into three categories: free, occupied, and unknown. Yamauchi [2] proposed to use *frontiers*, which are boundaries between free and unknown spaces, for viewpoint generation in exploration. Several criteria are possible in choosing one among frontiers such as closest to the robot [2] and maximizing the information gain [3]. The idea of frontier is very useful in viewpoint planning for exploration and has been expanded to, for example, multi-robot exploration [7] and/or multi-criteria exploration [8].

2.2 Coverage

When making an entire map of an environment, a robot needs to plan a set of viewpoints so that the whole environment is collectively observed. This is sometimes called as *coverage planning* and methods for Art Gallery Problem (AGP) [9] can be adopted. In robotics context, sensor coverage problems have been discussed (e.g., [10–12]). Ardiyanto and Miura [13] proposed a generalized coverage solver that can take into account of the cost of each viewpoint imposed by the environment and/or problem settings.

2.3 Object search

Ye and Tsotsos [5] proposed a framework for solving a visual object search problem. Using the probabilistic distribution of the target position and the probabilistic detection functions, the object search problem is formulated as a statistical optimization problem. Saidi et al. [14] proposed a similar approach to a 3D object search using a humanoid robot. Aydemir et al. [15] used domain knowledge on spatial relations between objects to guide a object search behavior. Masuzawa and Miura [4] treat an object verification problem as a viewpoint planning, which optimizes viewpoint sequences using a distance- and orientation-dependent object recognition model.

2.4 Integrated viewpoint planning

Multiple objectives are sometimes considered in exploration and/or mapping. Makarenko et al. [3] consider three utilities, information gain, navigability, and localizability, and choose a viewpoint which maximizes the total utility. Masuzawa and Miura [4] proposed a two-level hierarchical planner, in which the high-level part deals with determining the order of visits to viewpoint candidates, while the low-level part determines an actual viewpoint sequence to verify object candidates. They also introduced a loss function representing a deadline and showed that the robot's behavior changes for different loss functions. Diar and Miura [6] proposed a viewpoint planner which considers viewpoints for exploration and those of coverage in an integrated manner. However, they did not consider viewpoints for object verification.

3 Target Task and Overview of Integrated Viewpoint Planning

The task of the robot is to make a semantic map. A map describes not only 3D geometric information but also object types and locations. We use a realistic robot simulator V-REP [16] for experiments. Fig. 1 shows a simulated world and a mapping result. A simulated robot is equipped with an omnidirectional

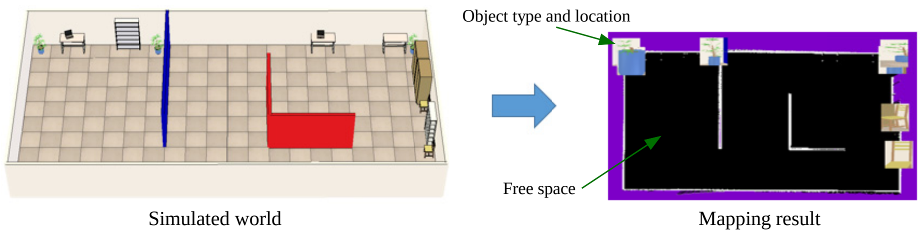


Fig. 1. Semantic mapping task.

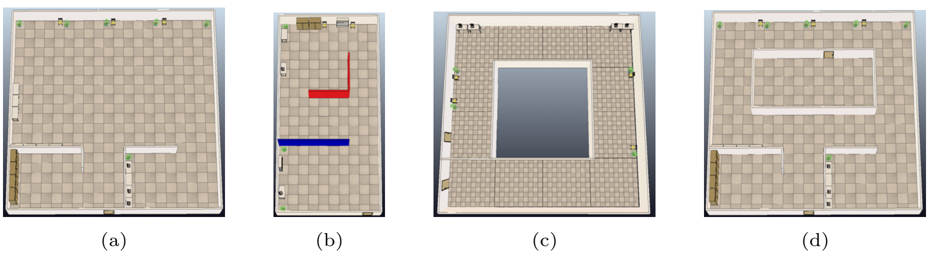


Fig. 2. Four environments used for experiments. From left to right, Environment A, B, C, and D. The size of Environment A, B, and D is $20m \times 20m$ while that of Environment B is $20m \times 20m$.

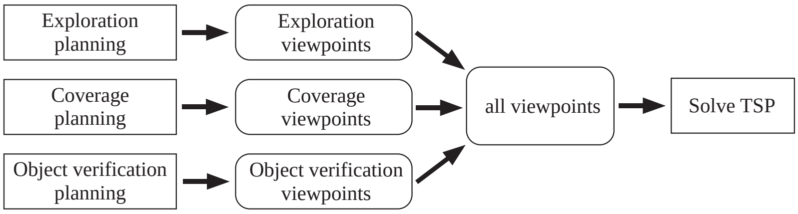


Fig. 3. Process of the integrated planning.

laser range finder (LRF) for free space recognition and an RGB-D camera for 3D measurement, object candidate detection, and object verification.

We consider the following three aspects in viewpoint planning for semantic mapping of unknown environments:

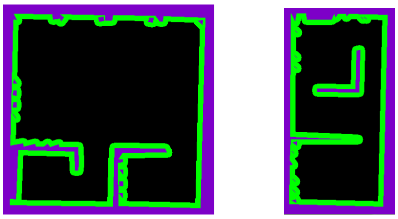
– *exploration planning* chooses viewpoints, measurements by the LRF from which can provide a large expansion of free spaces. Such free space information is necessary for the following two aspects of planning.
– *coverage planning* chooses viewpoints, observations by the camera from which can provide 3D space shape of the entire environment and object candidate information.
– *object verification planning* chooses viewpoints, observations by the camera from which can provide a more reliable information on the identity of object candidates.

Since the coverage planning is performed in already-explored regions and the object verification planning is performed for object candidates found in the coverage planning, one natural way to solve these planning problems is to perform them sequentially. That is, the robot first explores the whole environment and makes a 2D free space map. Then it observes the environment at planned viewpoints and makes a 3D description of the environment and enumerates detected object candidates. Finally it moves near to each object candidate and verify it. This sequential manner is obviously inefficient and an integrated approach is necessary.

Therefore we take an approach that the robot first generates viewpoints for these three planning aspects and chooses the best next viewpoint from a set of all generated viewpoints. The second step is done by solving a travelling salesman problem (TSP) and choosing the very first viewpoint. Fig. 3 illustrates the process of planning.

4 Integration of Exploration and Mapping Viewpoint Planning

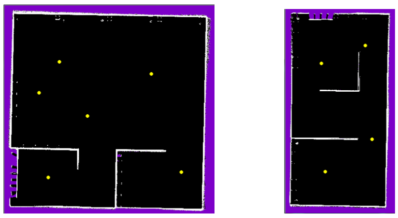
This section explains an integrated exploration and coverage planning by Sasongko and Miura [6]. The planning is composed of viewpoint generation for exploration, that for coverage, and integrated viewpoint selection.



(a) Informative regions of environment A.

(b) Informative regions of environment B.

Fig. 4. Example informative regions shown in green color.



(a) Optimal coverage viewpoints for environment A.

(b) Optimal coverage viewpoints for environment B.

Fig. 5. Optimal set of coverage viewpoints.

4.1 Viewpoint generation for exploration

We use a 2D occupancy grid map [17] to represent the free space of the environment. As a new LRF scan is input, a SLAM module updates the free space map. We adopt the frontier-based method [2]. A frontier point is a point which is in the free space region and adjacent to an unknown region. All of the frontier points are partitioned into clusters with a distance threshold th_d. We then calculate the centroid of each cluster of frontier points whose number of points is larger than a threshold th_n and determine the frontier point closest to the centroid as an exploration viewpoint. Finally we determine the exploration viewpoint V_e using the closest-frontier strategy [2]. We set the resolution of the grid map to $0.05m$ and use the thresholds $th_d = 0.25m$ and $th_n = 13$.

4.2 Viewpoint generation for coverage

Informative region In usual environments, objects are, for example, on tables and shelves and not on the floor. We would therefore like to limit the target regions for coverage to such *informative regions*. We currently use a simple heuristic by Okada and Miura [18]; that is, assuming that outline of 2D free spaces corresponds to such tables and shelves, the informative regions are defined as the ones which are within a certain distance to the free space boundaries. We currently set the distance to $1\,m$. Fig. 4 shows 2D maps and the corresponding informative regions for environments A and B shown in Fig. 2.

Coverage viewpoint generation We need to generate a set of viewpoints which covers all currently-known informative regions. The measurable range of the RGB-D camera and the incident angle limitation[1] are considered in coverage calculation. We adopt a visibility-based viewpoint planning [19]. This method

[1] The line-of-sight of the camera and the surface normal of an observed area should be within a certain angle. Currently, we use 80° as the threshold.

exploits a topological property of the free space, where coverage viewpoints can be put. A skeletonization technique is applied to the free space followed by a detection of junctions, which are then used as a set of viewpoint candidates V'_{c_can}. Since the skeletonization method is not always complete in a complex environment, we add a certain number of auxiliary viewpoint candidates V_{c_aux} to the set to get an updated set V_{c_can}, and then obtain an optimized set V_c by solving the Set Coverage Problem [13]. Fig. 5 shows examples of an optimized set of coverage viewpoints.

4.3 Integrated planning

The free space region expands as the robot moves and gets more data from the LRF. The robot calculates viewpoints and selects one of them at a certain timing, since it is not computationally efficient to update the set of viewpoints continuously and a frequent change of target viewpoint may make the robot move on a longer path. Therefore the robot generates a new set of the coverage viewpoints (V_c) for the currently-known informative regions, if the increment of the free space region is larger than a threshold th_f or if the entire environment has been already explored. Note that the exploration viewpoint (V_e) is updated every time as the robot gets data from the LRF. The value of th_f is determined experimentally as explained in Sec. 6.1.

We consider the union of V_e and V_c:

$$V = V_e \cup V_c$$

and choose the one from the set V. For this purpose, we solve the Travelling Salesman Problem (TSP) using 2-opt method [20], and the first viewpoint in the tour is used as the target viewpoint. The robot moves towards that viewpoint, and when the robot reaches there or the coverage viewpoint set is updated as mentioned above, a new TSP is solved.

5 Object Verification and Viewpoint Generation

5.1 Object candidate detection and verification

The robot searches the informative regions for objects. We use YOLOv2 [21] enhanced by COCO dataset [22] for object detection and verification. There are 80 classes for object candidates. YOLOv2 provides the confidence and a bounding frame of each detected object. We use the centroid of the frame as the detected position of the object. Fig. 6 shows an example object detection result.

We divide the detection results to high-confidence and low-confidence, using a threshold $th_o = 0.5$. Low-confidence objects must be verified by observing them again. In the case of Fig. 6, for example, the laptop and the vase are verified.

Fig. 6. YOLOv2 detection result. The confidence values of potted plant, vase, and laptop are 0.68, 0.35, and 0.27, respectively.

Table 1. Maximum verification distances for four objects.

No	Object Name	Object Picture	Max. Verification Dist.
1	Laptop		4m
2	Chair		4m
3	Potted plant		4m
4	Vase		3m

5.2 Verification viewpoint generation

In general, a closer observation increases the probability of correctly identifying an object, and such an idea should be considered in viewpoint generation for verification. Masuzawa and Miura [23] used a probabilistic observation model, which estimates the probability of successful verification from a set of observation parameters, for generating an observation plan. The model is for SIFT-based specific object recognition and was made for each object by actually observing the object from various viewpoints.

We take a similar empirical approach but use a much simpler model. That is, we assume that verification results depend only on the distance to each target object, and determine the maximum distance for which verification always succeeds. Table 1 summarizes the maximum verification distances for four objects used in this paper. The verification region of an object is thus defined as a circle centered at the object and with the radius of the maximum verification distance.

The object verification viewpoint (V_v) for an object is the point in the corresponding verification region and closest to the robot. If there are more than one object verification regions and they are overlapping, the object verification viewpoint is generated in the intersection region. Fig. 7 shows the verification viewpoint for an example scene shown in Fig. 6. Two among three objects are low-confidence ones. The intersection of the laptop and the vase verification re-

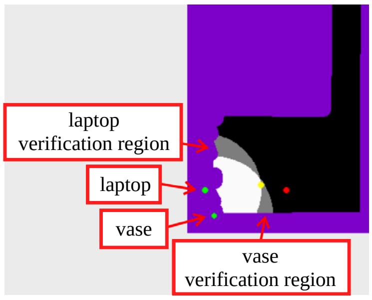


Fig. 7. Verification viewpoint for two objects. The white region is the intersection of two verification regions; light blue points are low-confidence objects and red point is the robot position. Yellow point is the chosen verification viewpoint.

gion is examined to choose the nearest position in it to the robot as the the verification viewpoint (shown in yellow).

5.3 Integrated planning

We consider the object verification viewpoint (V_v) in addition to the exploration and the coverage viewpoints (V_e and V_c, respectively). The method of viewpoint planning is similar to the one described before (see Sec. 4.3), that is, we consider the union of three types of viewpoints, V_e, V_c, and V_v:

$$V = V_e \cup V_c \cup V_v$$

and choose the one from the set V by solving the TSP (see Fig. 3).

The timing of updating viewpoints is slightly changed. In addition to the above-mentioned two conditions, we also consider if there is at least one newly detected low-confidence object.

6 Experimental Results

6.1 Determining the threshold for the increment of free space area

The proposed integrated planning method generates a new set of the coverage viewpoints (V_c) for the currently-known informative regions if the increment of the free space area is larger than a threshold th_f or the entire environment has been explored. If the threshold is too low, the robot updates V_c too often, thereby making the robot change the target subgoal frequently. If the threshold is too high, on the other hand, the robot rarely updates V_c, thereby making the robot ignore useful information of newly explored regions and newly detected objects. In either case, the robot behavior could be inefficient.

Using the integrated exploration and coverage planning explained in Sec. 4.3, we compared three values for th_f: $50m^2$, $100m^2$, and $150m^2$ in terms of the total

Table 2. Effect of th_f values on the total performance.

th_f	Environment A		Environment B	
	Time (min)	Distance (m)	Time (min)	Distance (m)
$50m^2$	11	72	9	79
$100m^2$	7	56	3.5	25
$150m^2$	9	64	5	34

time and the travelled distance for environments A and B. Table 2 summarizes the results, showing a tendency that too low and too high th_f values are not desirable. From the results, we decided to use $th_f = 100m^2$.

6.2 Mapping results and comparison among strategies

We show the results of the following two strategies:

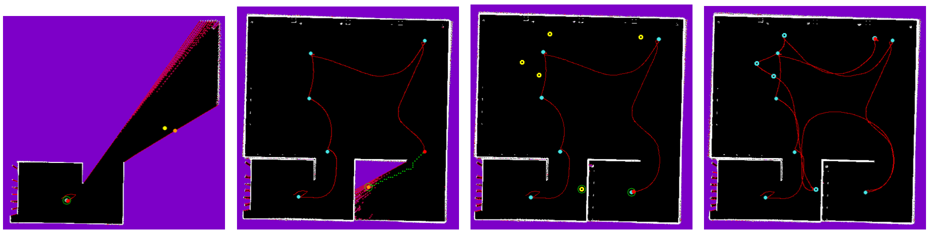
– Sequential strategy (ST_{seq}) which performs a geometric mapping (i.e., exploration and coverage) and a semantic one (i.e., object verification and localization) sequentially.
– Integrated strategy (ST_{int}) which perform both the geometric and the semantic mapping in an integrated manner. This is the proposed strategy.

Figs. 8 and 9 show the mapping process of the sequential and the integrated strategy for Environment A and D shown in Fig. 2. In the sequential strategy, the robot first explores the environment to make a geometric map, followed by another round of navigation for verifying low-confidence objects found in the way of geometric mapping. As a result, there are redundant movements in many places in the environments.

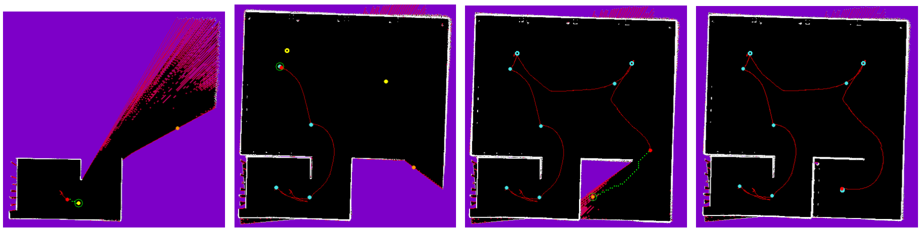
Table 3 summarizes the quantitative data for the strategies. The table also includes the results of geometric mapping (not including object detection and verification) by two strategies for comparison; one strategy is sequential exploration and coverage (ST^g_{seq}) and the other is the integrated exploration and coverage (ST^g_{int}). Comparison results for strategies ST_{seq} and ST_{int} show the effectiveness of integrated viewpoint planning. Note that strategy ST_{int} performs extra object recognition tasks compared to those for geometric mapping, but requires a shorter traveling distance than their sequential version (ST^g_{seq}).

7 Conclusions and Future Work

This paper has presented an integrated viewpoint planning for semantic mapping of unknown environments, which includes exploration, coverage for 3D mapping, and object detection and localization. Since generating the optimal sequence of viewpoints is hard without any prior information of the environment, we repeatedly generate a locally-optimal viewpoints candidates based on the information



Mapping process by Sequential Strategy (ST_{seq})



Mapping process by Integrated Strategy (ST_{int})

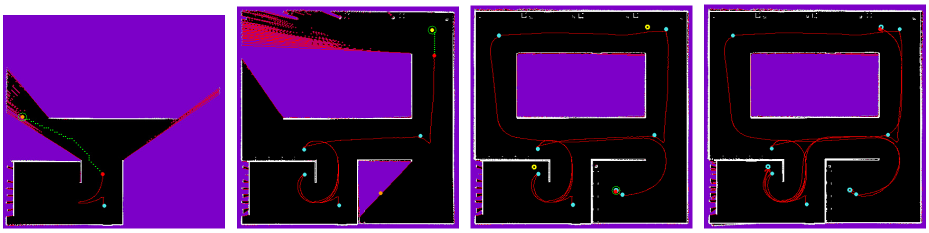
Fig. 8. Mapping process of the sequential and the integrated strategy for Environment A. Red lines indicate the robot path; blue points indicate visited viewpoints; yellow points indicate viewpoint candidates and those with green circles indicate the current target viewpoint.

Table 3. Quantitative comparison of four strategies.

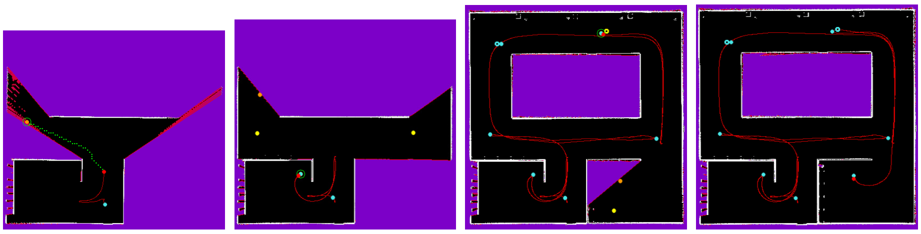
Strategy	Environment A		Environment B		Environment C		Environment D	
	Time (min)	Distance (m)	Time (min)	Distance (m)	Time (min)	Distance (m)	Time (min)	Distance (m)
ST_{seq}	32	113.8	22	70.7	16	80.9	26	146.6
ST_{int}	13	64.2	10	45.1	9	50.9	20	105.5
ST^g_{seq}	14	77.4	8	47.3	13	88.4	15	121.8
ST^g_{int}	7	56.0	3.5	25.0	7	50.5	9.5	98.4

of newly explored regions and newly detected object candidates. We experimentally determine a good interval for viewpoint updates. We tested the proposed algorithm in a realistic robotic simulation environment to show the efficiency of our integrated planning strategy.

Implementing the method on a real robot for evaluating in real situations is future work. To do this, the object recognition part needs to be enhanced largely, especially in the observation models. The current model considers only the distance to an object but a more variety of factors such as orientation and lighting conditions should also be taken into account.



Mapping process by Sequential Strategy (ST_{seq})



Mapping process by Integrated Strategy (ST_{int})

Fig. 9. Mapping process of the sequential and the integrated strategy for Environment D. Red lines indicate the robot path; blue points indicate visited viewpoints; yellow points indicate viewpoint candidates and those with green circles indicate the current target viewpoint.

Acknowledgment

This work is in part supported by JSPS KAKENHI Grant Number 17H01799 and the Hibi Science Foundation.

References

1. S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. The MIT Press, 2005.
2. B. Yamauchi. A Frontier-Based Approach for Autonomous Navigation. In *Proceedings of the 1997 IEEE Int. Conf. on Computational Intelligence in Robotics and Automation*, pp. 146–151, 1997.
3. A.A. Makarenko, S.B. Williams, F. Bourgault, and H.F. Durrant-Whyte. An Experiment in Integrated Exploration. In *Proceedings of 2002 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 534–539, 2002.
4. H. Masuzawa and J. Miura. Observation Planning for Environment Information Summarization with Deadlines. In *Proceedings of 2010 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 30–36, 2010.
5. Y. Ye and J.K. Tsotsos. Sensor Planning for 3D Object Search. *Computer Vision and Image Understanding*, Vol. 73, No. 2, pp. 145–168, 1999.
6. D.F. Sasongko and J. Miura. An Integrated Exploration and Observation Planning for an Efficient Indoor 3D Mapping. In *Proceedings of 2017 Int. Conf. on Mechatronics and Automation*, pp. 1924–1929, 2017.

7. J. Faigl and M. Kulich. On Benchmarking of Frontier-based Multi-robot exploration strategies. In *Proceedings of 2015 European Conf. on Mobile Robots*, pp. 1–8, 2015.
8. N. Basilico and F. Amigoni. Exploration Strategies Based on Multi-Criteria Decision Making for Search and Rescue Autonomous Robots. *Autonomous Robots*, Vol. 31, No. 4, pp. 401–417, 2011.
9. J. O'Rourke. *Art Gallery Theorems and Algorithms.* Oxford University Press, 1987.
10. H. González-Baños and J.-C. Latombe. A Randomized Art-Gallery Algorithm for Sensor Placement. In *Proceedings of 7th Annual Symp. on Computational Geometry*, pp. 232–240, 2001.
11. P.K. Agarwal, E. Ezra, and S.K. Ganjugunte. Efficient Sensor Placement for Surveillance Problems. In *Proceedings of the 5th IEEE Int. Conf. on Distributed Computing in Sensor Systems*, pp. 301–314, 2009.
12. V. Ramaswamy and J.R. Marden. A Sensor Coverage Game with Improved Efficiency Guarantees. In *Proceedings of 2016 American Control Conf.*, pp. 6399–6404, 2016.
13. I. Ardiyanto and J. Miura. Generalized Coverage Solver Using Hybrid Evolutionary Optimization. *Int. J. of Innovative Computing, Information and Control*, Vol. 13, No. 3, pp. 921–940, 2017.
14. F. Saidi, O. Stasse, K. Yokoi, and F. Kanehiro. Online Object Search with a Humanoid Robot. In *Proceedings of 2007 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 1677–1682, 2007.
15. A. Aydemir, K. Sjöö, J. Folkesson, A. Pronobis, and P. Jensfelt. Search in the Real World: Active Visual Object Search Based on Spatial Relations. In *Proceedings of 2011 IEEE Int. Conf. on Robotics and Automation*, pp. 2818–2824, 2011.
16. E. Rohemr, S.P. Singh, and M. Freese. V-REP: A Versatile and Scalable Robot Simulation Framework. In *Proceedings of 2013 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 1321–1326, 2013.
17. A. Elfes. Sonar-Based Real-World Mapping and Navigation. *Int. J. of Robotics and Automat.*, Vol. 3, No. 3, pp. 249–265, 1987.
18. Y. Okada and J. Miura. Exploration and Observation Planning for 3D Indoor Mapping. In *Proceedings of 2015 IEEE/SICE Int. Symp. on System Integration*, pp. 599–604, 2015.
19. I. Ardiyanto and J. Miura. Visibility-based Viewpoint Planning for Guard Robot using Skeletonization and Geodesic Motion Model. In *Proceedings of the 2013 IEEE Int. Conf. on Robotics and Automation*, pp. 652–658, 2013.
20. O. Mersmann, B. Bischl, J. Bossek, H. Trautmann, M. Wagner, and F. Neumann. Local Search and the Traveling Salesman Problem: A Feature-Based Characterization of Problem Hardness. In *Learning and Intelligent Optimization*, Vol. LNCS 7219, pp. 115–129. 2012.
21. J. Redmon and A. Farhadi. YOLO9000: Better, Faster, Stronger. arXiv:1612.08242, 2016.
22. S. Belongie L. Bourdev R. Girshick J. Hays P. Perona D. Ramanan C. L. Zitnick T.-Y. Lin, M. Maire and P. Dollar. Microsoft COCO: Common Objects in Context. arXiv:1405.0312, 2014.
23. H. Masuzawa and J. Miura. Observation Planning for Efficient Environment Information Summarization. In *Proceedings of 2009 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 5794–5780, 2009.