The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Bayesian Robot Localization using Spatial Object Contexts

Chuho Yi, Il Hong Suh, Gi Hyun Lim, and Byung-Uk Choi

Abstract— **We propose a semantic representation and Bayesian model for robot localization using spatial relations among objects that can be created by a single consumer-grade camera and odometry. We first suggest a semantic representation to be shared by human and robot. This representation consists of perceived objects and their spatial relationships, and a qualitatively defined odometry-based metric distance. We refer to this as a topological-semantic distance map. To support our semantic representation, we develop a Bayesian model for localization that enables the location of a robot to be estimated sufficiently well to navigate in an indoor environment. Extensive localization experiments in an indoor environment show that our Bayesian localization technique using a topological-semantic distance map is valid in the sense that localization accuracy improves whenever objects and their spatial relationships are detected and instantiated.**

I. INTRODUCTION

Service robots perform complex tasks whose execution requires frequent interaction with humans. Imagine visiting a university and asking a security guard how to get to Professor Suh's office. The guard might say, "Go to the IT building and proceed along the corridor to your right, where you will see restrooms on your left as well as classrooms. When you reach the end of the corridor, turn left. At that point, you will find the office on your right." You would find the correct room without any difficulty following these instructions. Humans do not necessarily use accurate quantitative information to perceive space in the current location or for traveling to another location. Instead, humans remember a few landmarks that define the space. Based on a specific structure or distinct objects, humans restructure their knowledge based on spatial contexts and then reuse the knowledge [15]. This method may not define the exact location quantitatively, but as many fragments of spatial context accumulate, it enables a sufficiently high level of space recognition and localization.

It is necessary to build an accurate metric map and a semantic map for symbolic inference for the current robot localization problem. This requires symbol grounding between the metric data and the semantic representation. An accurate sensor is essential for the execution of complicated maneuvers. In the semantic map proposed by Kuipers et al. [6], the semantic structure of space is inferred using the model built, based on a hierarchy of successive environmental data accumulated while the robot moves around. In addition, a global map is derived by integrating the global topology information and the local metric data.

Many robot localization methods have been studied over the last decade, using grid-based maps [5], feature-based maps [8][20], topological maps [7], semantic maps [6], or selecting diverse types of maps adaptively [9].

In research on the probabilistic Bayesian model, Thrun et al. applied Bayes' rule to robot localization [3]. The topological map is represented in semantics using ontology, and relies on semantics to infer navigation [10][11].

In a recent semantic approach, Vasudevan et al. proposed a hybrid map that relies on numerical data in an indoor environment and uses topological information for the integration of each space [2][16]. Based on this approach, we constructed a probabilistic object graph using the spaces of recognized objects. In addition, we propose place classification using such a representation.

Ranganathan et al. proposed a method of semantic location modeling for space recognition using the semantic information of the recognized objects [1]. They suggested a Bayesian inference structure that contains probabilistic models for low-level features up to high-level inference. There, the increase in computational complexity was resolved with a Markov chain Monte Carlo approach.

Semantic representation is necessary for human-interacting service robots to take orders and complete tasks. Until now, we have concentrated on robot-centered knowledge that enables humans to interact with robots [13][17].

Most research on high-level knowledge has detected objects in camera images and estimated the distance from the camera to the observed objects to create relationships among the objects. However, these quantities are inaccurate due to lens distortion and incorrect feature detection and matching. This leads to errors when applied to real environments.

To cope with these issues, we suggest a semantic representation and Bayesian model for spatial relationships among objects, and show that our representation is useful for the localization of a mobile robot. The spatial context used in the proposed semantic representation includes the observed objects, the robot-to-object (r-o) distance context that represents the distance from the robot to a particular object, the r-o bearing context that represents the direction from the robot to a particular object, and the object-to-object (o-o) relationship context that denotes the relationship among objects. For this, we use visual pattern recognition to

This work was supported by the Technology Innovation Program of MKE/KEIT [2008-F-038-01, Development of Context Adaptive Cognition Technology]

Chuho Yi is with the Division of Electrical and Computer Engineering, Hanyang University, Korea d1uck@hanyang.ac.kr

Il Hong Suh is with the College of Information and Communications, Hanyang University, Korea. All correspondences should be addressed to I. H. Suh ihsuh@hanyang.ac.kr

Gi Hyun Lim is with the Division of Electrical and Computer Engineering, Hanyang University, Korea hmetal@hanyang.ac.kr

Byung-Uk Choi is with the Division of Electrical and Computer Engineering, Hanyang University, Korea buchoi@hanyang.ac.kr

recognize objects and estimate rough metric data [18].

We next propose a topological-semantic distance map consisting of spatial object contexts and spatial robot contexts. In a topological-semantic distance map, a node is one of the components of a general topological map that plays the role of a standard and contains information on the spatial object contexts. In addition, spatial robot contexts used in the proposed semantic representation can explain an approximate distance and bearing from one assigned node to another. We describe how an approximate qualitative distance is the node-to-node (n-n) distance context and qualitative bearing is the n-n bearing context. Global localization is therefore made possible by the object in the topological-semantic distance map and the spatial object contexts. In addition, the localization is processed more specifically and locally based on the information gathered from the global localization, deriving the probability distribution of a robot possibly locating itself around the node. Therefore, we propose the probabilistic Bayesian model by which localization accuracy can be improved as many fragments of spatial object contexts are continually obtained, and eventually verify the practicality of the proposed methods through an experiment so that robot localization is sufficiently qualified for navigation in indoor environments. One of the most important contributions of the paper is the manner in which noise in the measurements provided by the ERSP(Evolutionary Robotics Software Package) system with a single camera is reasonably managed.

II. Topological-Semantic Distance Map

Almost all data in our methods are represented semantically by means of ontology, which ensures that only sound and complete data are asserted and propagated with ontology inference. Noisy sensor data such as false-positives and true-negatives can be filtered using the relationships and rules of logical reasoning. In many cases of false-positive, the properties are illogical, such as when a misclassified object is floating in the air or penetrating walls or other objects. These cases can be evaluated logically by axiomatic rules, and the robot will know what to expect in the next step. That enables robots to predict and pay attention a priori.

A. Spatial Object Contexts

Figure 1 shows the changes in real relationships between a robot and observed objects according to the robot's location transition from x to x' in the real world. Parameters r, ζ, and ω are the metric distance of the object relative to the robot, the metric bearing of the object relative to the robot, and the bearing among objects in real-world coordinates, respectively. The subscripts indicate the indexes of the objects observed.

Figure 2 shows the changes in metric relationships between a robot and observed objects according to the robot location transition from x to x'. Parameters $\tilde{r}$, $\tilde{\omega}$, and $\tilde{\zeta}$ denote the estimated metric distance of an object relative to the robot, the metric bearing of an object relative to the robot, and the bearings among objects in the robot coordinates, respectively, measured by a single camera. In general, the

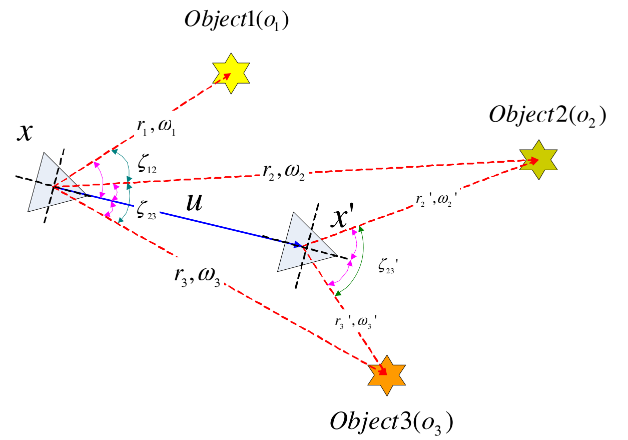


Fig. 1. Spatial relationships between robot and objects in the real world.

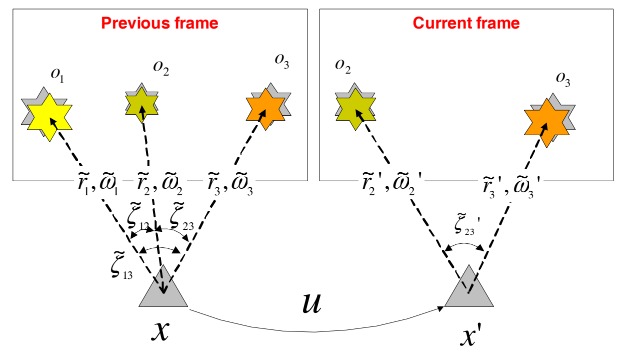


Fig. 2. Metric relationships between robot and objects in image sequences.

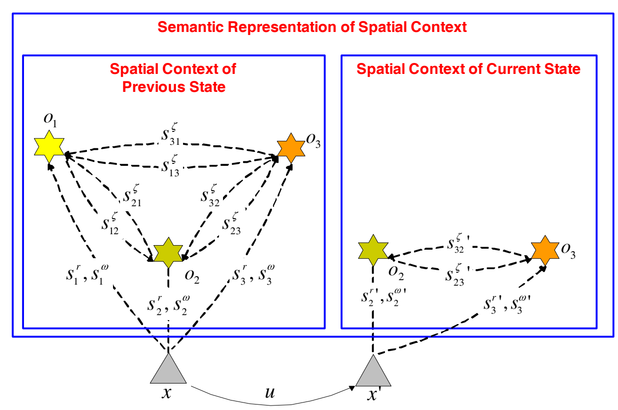


Fig. 3. Semantic representation comprising a set of respective spatial object contexts among the observed objects and the robot.

distance and orientation obtained using a commercial vision system such as the ERSP vision system with a single camera are often inaccurate. In cases when a single camera is attached to a robot, the metric data quantities such as distance and orientation to an object from the robot become too unreliable to be used. Thus, for semantic representation, the metric data are linked to the symbols of the spatial relationships, according to the given conditions. However, we use the information that mutual association exploits the geometric relationship between objects (landmarks) [19].

TABLE I
SEMANTIC REPRESENTATION INCLUDING ALL OF THE SPATIAL OBJECT CONTEXTS IN FIGURE 3.

State	Semantic Representation
Previous state	nearby(o_1, Robot), left front(o_1, Robot), right near(o_1, o_2), right far(o_1, o_3), far(o_2, Robot), front(o_2, Robot), left near(o_2, o_1), right near(o_2, o_3), far(o_3, Robot), right front(o_3, Robot), left far(o_3, o_1), left near(o_3, o_2)
Current state	near(o_2, Robot), left front(o_2,Robot), right far(o_2, o_3) nearby(o_3, Robot), right front(o_3, Robot) left far(o_3, o_2)

Object recognition is a fundamental factor in semantic representation. In general, an object is recognized visually by measuring the similarity between features of observed objects and those of corresponding object models. In this section, we use scale-invariant feature transform (SIFT) features that are known to be invariant to scale and rotation changes [4][12].

The distance from the robot to the observed object is estimated with a single camera to derive a fragment of spatial context. After an object is recognized, its height in the image space is measured using a set of corresponding features, and then a metric distance is estimated by ERSP [18].

B. Semantic Representation of Spatial Object Contexts

Figure 3 shows a semantic representation consisting of observed objects and their respective spatial symbols. Here, the spatial context includes distance, bearing, and relationship contexts. The r-o distance context denoted by s^r is the distance of the object from the robot. Each distance context is represented by one of a set of distance symbols, that is, $s^r = \{nearby, near, far\}$. The r-o bearing context denoted by $s^\omega = \{$ *front*, *left front*, *left*, *left rear*, *rear*, *right rear*, *right*, *right front*$\}$ is the bearing of the object relative to the robot. The o-o relationship context denoted by $s^\zeta = \{$ *left far*, *left near*, *left nearby*, *right nearby*, *right near*, *right far*$\}$ is the relationship among objects.

Table 1 shows a semantic representation using symbols for all of the spatial contexts in Figure 3. Our robot localization application finds the position of the robot using only these types of semantic representations with qualitative metric data.

C. Topological-Semantic Distance Map

Neither the method we propose here nor the experiments we conducted focused on topological map-building, which can be considered merely a placeholder scheme from the robotic research point of view.

Figure 4 shows an example of topological nodes obtained using a laser sensor. In our experiments, we registered topological nodes when the robot was located at positions defined by a generalized Voronoi graph. Topological-semantic distance-map building takes place in three steps:

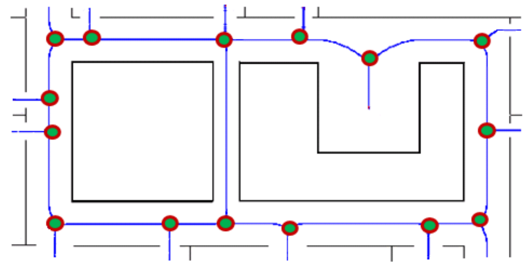


Fig. 4. Example of topological nodes using a generalized Voronoi graph of the 4th floor of the IT building at Hanyang University, Seoul, Korea.

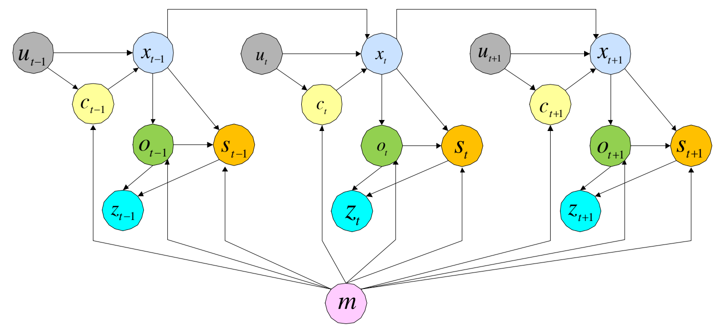


Fig. 6. Graphical model of Bayesian localization.

1) Topological nodes from a range sensor are assigned to a distinct space.
2) Observed objects and estimated spatial object contexts are registered in the topological-semantic distance map as the robot rotates 360° around a fixed position.
3) The distance context and the bearing context estimated during robot movement is added between nodes. This is considered a spatial robot context of n-n contexts.

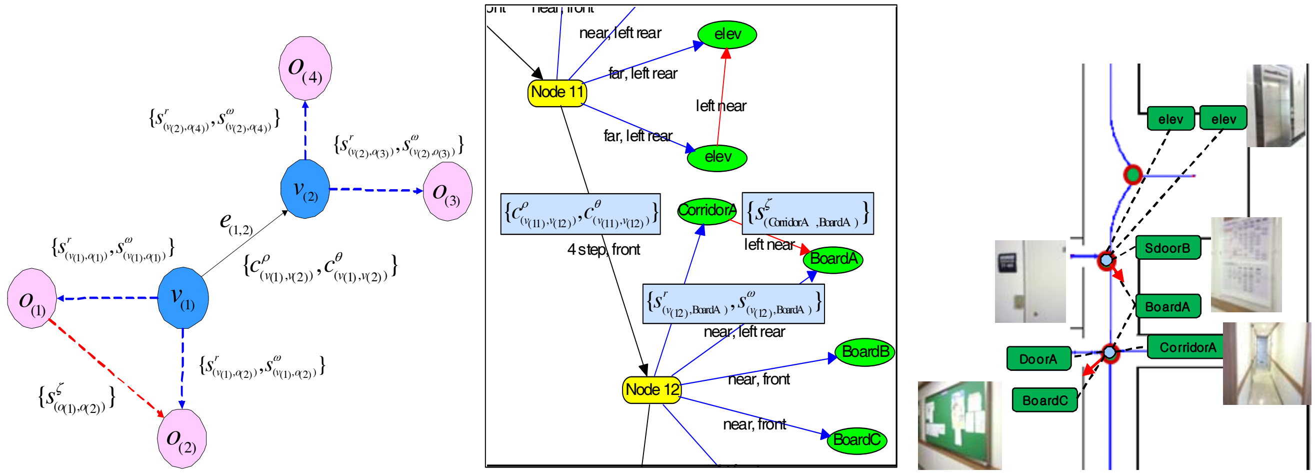
The first step is identical to the assignment of distinct nodes. We used a laser sensor (SICK 5000) to assign nodes that were subsequently processed to build a topological map [7]. In the second step, observed objects and spatial object contexts were added to the topological-semantic distance map. Each spatial object context individually determines for which current robot it locates at node. In the third step, the spatial robot contexts estimated by odometry between nodes are added to the topological-semantic distance map. These contexts constitute relative information between nodes. Figure 5 gives an example of a topological-semantic distance map.

III. BAYESIAN MODEL FOR SEMANTIC LOCALIZATION

In this section, we describe our Bayesian models and their mathematical formulations. We rely partially on the probabilistic Bayesian model proposed by Thrun et al.[3] and Ranganathan et al. [1] to formulate our probabilistic models.

A. Probabilistic Localization Model

Figure 6 depicts a graphical model of Bayesian localization. The robot is given a map and its goal is to determine its location relative to this map given observations of the environment and the robot's movements.



(a) (b) (c)

Fig. 5. Example of a topological-semantic distance map. (a) Semantic representation comprising a set of respective spatial object contexts and spatial robot contexts. (b) Specific example of a topological-semantic distance map comprising a set of respective spatial object contexts and spatial robot contexts. (c) Relationship between topological-semantic distance map and the real environment.

In the figure, robot location is denoted by $x = [v_{(i)}\ e_{(i,j)}\ \rho_{(i,j)}\ \theta_{(i,j)}]^T$ where $v_{(i)}$ and $e_{(i,j)}$ are the topological node and edge, respectively. In addition $\rho_{(i,j)}$ and $\theta_{(i,j)}$ are metric distance and bearing on edge $e_{(i,j)}$, respectively. The object is denoted by o. This metric-topological framework enables localization to be bounded globally, the map size to increase monotonically in dimensionality, and the location to be calculated locally between two nodes. A set of semantics for spatial object contexts containing distance, bearing, and relationship is denoted by $s = \{s^r, s^\omega, s^\zeta\}$. A set of semantics for spatial robot contexts containing distance and bearing is denoted by $c = \{c^\rho, c^\theta\}$. We represent the features extracted from the image with z and the map is denoted by m. We have already discussed the type of map, the topological-semantic distance map.

Probabilistic robot location represents beliefs through conditional probability distribution. We will denote belief over a state x_t by $Bel(x_t)$, which is an abbreviation for the following:

$$Bel(x_t) = p(x_{0:t}|s_{1:t}, o_{1:t}, z_{1:t}, c_{1:t}, u_{1:t}, m) \tag{1}$$

Psterior $Bel(x_t)$ is obtained in a similar way as the derivation using Bayes' rule and the Markov assumption. In particular, we have

$$\begin{aligned} &Bel(x_t) \\ &= \eta \cdot p(s_t, o_t, z_t|x_t, m) \int p(x_t|x_{t-1}, c_t, u_t) Bel(x_{t-1}) dx_{t-1} \end{aligned} \tag{2}$$

where the probability $p(s_t, o_t, z_t|x_t, m)$ is the semantic measurement, the probability $p(x_t|x_{t-1}, c_t, u_t)$ is the state transition, and $Bel(x_{t-1})$ is belief at time $t-1$, respectively. The prediction model is $\int p(x_t|x_{t-1}, c_t, u_t) Bel(x_{t-1}) dx_{t-1}$. The probabilistic localization model is divided into two parts, namely, the measurement model and the prediction model, which correspond to two terms on the right side of Eq. (2). The measurement model uses the semantic representation and is the main focus of this work. Contexts are uncertain data so they should be approximated with a stochastic distribution. The focus of this section is the measurement model and contexts of objects. Here, we assume that sensors are uncertain, and a lower distribution is approximated because fewer semantics are available. The reverse would also be true. We calculate the location posterior using the semantic representation described in Section II. B.

B. Semantic Measurement Model

The semantic measurement model is based on the joint probability of robot location x, map m, spatial object contexts s, object o, and extracted feature z, From the graphical model in Figure 6, the joint probability can be written as

$$p(s_t, o_t, z_t|x_t, m) = \frac{p(x_t)p(o_t|x_t, m)p(s_t|x_t, o_t, m)p(z_t|s_t, o_t)}{p(x_t, m)} \tag{3}$$

Here, we assume that the probability $p(x_t)$ is the same as $p(x_t, m)$ because the robot position x_t is located on map m. Applying Bayes' law to the semantic measurement model from Eq. (3) gives

$$\begin{aligned} &p(s_t, o_t, z_t|x_t, m) \\ &\quad = p(o_t|x_t, m)p(s_t|x_t, o_t, m)\tfrac{p(z_t)p(o_t|z_t)p(s_t|o_t, z_t)}{p(o_t)p(s_t|o_t)} \end{aligned} \tag{4}$$

We assume that the probability $p(z_t)$ has a uniform distribution. Therefore, we can obtain the distribution of the measurement model as

$$\begin{aligned} &p(s_t, o_t, z_t|x_t, m) \\ &\quad = \eta \cdot p(o_t|z_t)p(s_t|o_t, z_t)p(o_t|x_t, m)p(s_t|x_t, o_t, m) \end{aligned} \tag{5}$$

where η is the normalization constant and $p(o_t|z_t)$ is a term related to object recognition. It is evaluated based on the similarity between extracted features of the observed objects and the features of the corresponding object models.

We use a supervised approach to building the object models. Training data consist of images, each of which contains only one object. These images are captured at every known reference distance. The object model consists of features for object recognition, distance from the camera to the corresponding object, and the height of the object in pixels.

The $p(s_t|o_t, z_t)$ is the likelihood of similarity in spatial context between the estimated metric data for the observed object and the spatial contexts. It is evaluated based on the similarity between extracted features of the observed objects and the features of the corresponding object models.

Spatial contexts of objects are computed from the results of estimated metric distances and bearings as

$$p(s_t|o_t, z_t) = \{p(s_t^d|o_t, z_t), p(s_t^b|o_t, z_t), p(s_t^r|o_t, z_t)\} \quad (6)$$

$$p(s_t^r|o_t, z_t) = \mathbf{N}(r^s - \tilde{r}, {\sigma_r}^2) \quad (7)$$

$$p(s_t^\omega|o_t, z_t) = \mathbf{N}(\omega^s - \tilde{\omega}, {\sigma_\omega}^2) \quad (8)$$

$$p(s_t^\zeta|o_t, z_t) = \mathbf{N}(\zeta^s - \tilde{\zeta}, {\sigma_\zeta}^2) \quad (9)$$

where $\mathbf{N}(\mu, \sigma^2)$ is normally distributed with mean μ and variance σ^2. In addition, r^s, ω^s, and ζ^s are trained at every known reference for the spatial context of distance, bearing and relationship, respectively. The variances of σ_r^2, σ_ω^2, and σ_ζ^2 are introduced to reflect the uncertainty of vision sensor. The further an object is from the robot or from other objects, the more inaccurate the metric distance and bearing will be. More specifically, dividing the spatial context more finely will improve localization performance.

The probability $p(o_t|x_t, m)$ is the likelihood of similarity between the observed objects in the current state and those in the previous state, and formulated as

$$p(o_t|x_t, m) = exp(-\|o - o_x\|^2) \quad (10)$$

where o and o_x represent the observed object in the current and previous states, respectively. The probability $p(s_t|x_t, o_t, m)$ is the likelihood of the spatial context.

The likelihood of the spatial context is computed as

$$\begin{aligned} &p(s_t|x_t, o_t, m) \\ &= \textstyle\sum_a^N [f_c(s^{r_a} - s_x^r) f_c(s^{\omega_a} - s_x^\omega) \sum_b^N f_c(s^{\zeta_{ab}} - s_x^\zeta)] \end{aligned} \quad (11)$$

where $f_c(s_a, s_x) = exp(-\|s_a - s_x\|^2)$, and s and s_x represent the estimated spatial contexts of an object in the current and previous states, respectively. Each spatial context of an object belongs to one of some number of different distributions. Each context in Eq. (11) is described by a component probability density function, and its mixture of distributions is the probability that an observation comes from this component. Here, a mixture of three normal distributions with different means may result in a density with three spatial contexts of the object, which is not modeled by standard parametric distributions.

C. Prediction Model and Recursive Bayesian Model

The last term on the right side of Eq. (2) is an update term. The control model u represents simple motion data related to the state transition, as follows:

$$\begin{bmatrix} v'_{(i)} \\ e'_{(i,j)} \\ \rho'_{(i,j)} \\ \theta'_{(i,j)} \end{bmatrix} = \begin{bmatrix} v_{(i)} \\ e_{(i,j)} \\ \rho_{(i,j)} \\ \theta_{(i,j)} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \Delta\rho \cdot \cos(\theta_{(i,j)} + \Delta\theta - \theta^c_{(i,j)}) \\ \Delta\theta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \varepsilon_\rho \\ \varepsilon_\theta \end{bmatrix} \quad (12)$$

where $\Delta\rho$ and $\Delta\theta$ are the robot motion of bearing and movement, respectively, while ε_ρ and ε_θ are a zero-mean error variable of motion and bearing.

The localization posterior is calculated using Eq. (2), however, the modification is not applied directly to our Bayesian model because Eq. (2) is not a predictable term. To solve this problem, we modify Eq. (2) in the localization model to a particle filter, a type of recursive Bayesian estimation [3][14], to manage the complicated computations.

In particle filters, the samples of a posterior distribution are called particles and are denoted by

$$\chi_t := x_t^{[1]}, x_t^{[2]}, ..., x_t^{[Y]} \quad (13)$$

Each particle $x_t^{[y]}$ (with $1 \le y \le Y$) is a concrete instantiation of the state at time t. So-called importance factors are used for each particle $x_t^{[y]}$ to incorporate the measurement into the particle set. The importance is thus the probability of the measurement under the particle, given by

$$\begin{aligned} w_t^{[y]} &= \frac{\text{target distribution}}{\text{proposal distribution}} \\ &= p(o_t|z_t) p(s_t|o_t, z_t) p(o_t|x_t^{[y]}, m) p(s_t|x_t^{[y]}, o_t, m) \end{aligned} \quad (14)$$

By resampling particles with a probability proportional to $w_t^{[y]}$, the resulting particles are indeed distributed according to the product of the proposal and the importance weights $w_t^{[y]}$.

$$Bel(x_t) = \eta' \cdot w_t^{[y]} \sum_y^Y p(x_t^{[y]}|x_{t-1}^{[y]}, c_t, u_t, m) Bel(x_{t-1}) \quad (15)$$

Figure 7 shows a large particle distribution that denotes the initial robot location, and then becomes smaller with recursion, reflecting a number of spatial object contexts.

Fig. 7. Example of a concentrated distribution with recursion reflecting a number of contexts.

Fig. 8. Examples of trained object(landmark) images.

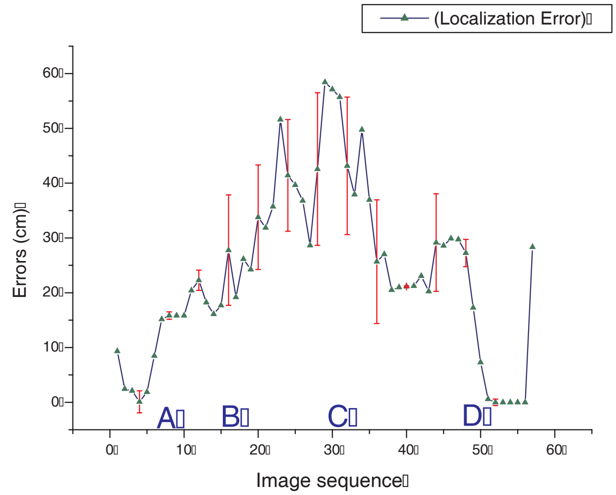


Fig. 11. Localization errors in our experiment using spatial object contexts.

IV. Experimental Results

A Pioneer 3 AT robot carrying a single consumer-grade camera was driven around an indoor environment (14 $\times$ 26.5 m) to evaluate the performance of the proposed localization process.

Figure 8 shows examples of trained object images selected by the user. The camera observed 16 objects during its travels. Distinctive objects such as doorplates, bulletin boards, and panel boards were used for object recognition.

Figure 9 illustrates the topological-semantic distance map consisting of 15 nodes (yellow, rectangle) and 42 objects (green, circle). Solid lines between nodes are the edges that represent n-n contexts of distances and bearings. The dotted blue lines denote the r-o context and the dotted red lines represent the o-o context. The topological-semantic distance map is included as an ontological representation for robot knowledge. Figure 10 shows part of the ontology schema and map instance.

The experiment showed an 88.93% probability of exact accuracy in topological localization. However, 13.07% of the probable error only occurs around the neighboring node. The reason that this error occurs is that a transition is processed between a start node to a goal node in the transition between edges.

Figure 11 shows the localization errors from node 11 to node 12 in the proposed semantic localization. Even though the robot's location was initially estimated accurately, the errors increased due to the lack of semantics obtained in subsequent movements. When objects far from the robot were observed, the localization error and standard deviation were relatively high. However, the distribution of robot locations became smaller as the robot continued to navigate and the observed objects became closer. If the previously observed object disappeared due to the rotation of the robot, the standard deviation increased. In this experiment, the maximum error was 68.04 cm, and the mean and standard deviation were 27.3 and 21.29 cm, respectively.

At point A in Figure 11, the distribution of robot locations became smaller due to observations of the objects near the robot. At point B, the object disappeared due to the rotation of the robot, thereby spreading out the distribution. At point C, the observed objects were far from the robot, leading to high uncertainty in the spatial context. The error in robot location as well as the distribution became larger. However, the situation at point D shows that robot location error and its corresponding distribution decreased as the robot moved closer to objects.

V. Conclusions

This paper proposed a semantic representation and Bayesian model for robot localization using spatial contexts among objects, and described them using symbols. Our proposed method enables robots to be localized using spatial object contexts and their probabilistic models. Experimental results of our proposed Bayesian robot localization scheme in an indoor environment demonstrated that as contextual evidence increased, location accuracy improved despite using an inaccurate sensor such as a consumer-grade camera.

References

[1] A. Ranganathan and F. Dellaert, "Semantic Modeling of Places using Object," *Robotics: Science and Systems (RSS)*, 2007.

[2] S. Vasudevan, V. Nguyen, and R. Siegwart, "Cognitive Maps for Mobile Robots - An Object based Approach," *Proceedings of the IROS*, pp 7-12, 2006.

[3] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*, MIT Press, Cambridge, MA, 2005.

[4] D.G. Lowe, "Distinctive image features from scale invariant keypoints," *Int'l Journal of Computer Vision, Vol. 60*, no 2, pp. 91-110, 2004.

[5] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *IEEE Computer, Vol.22*, pp.44-57, 1989.

[6] B. Kuipers, "The Spatial Semantic Hierarchy," *Artificial Intelligence*, pp.191-233, 2000.

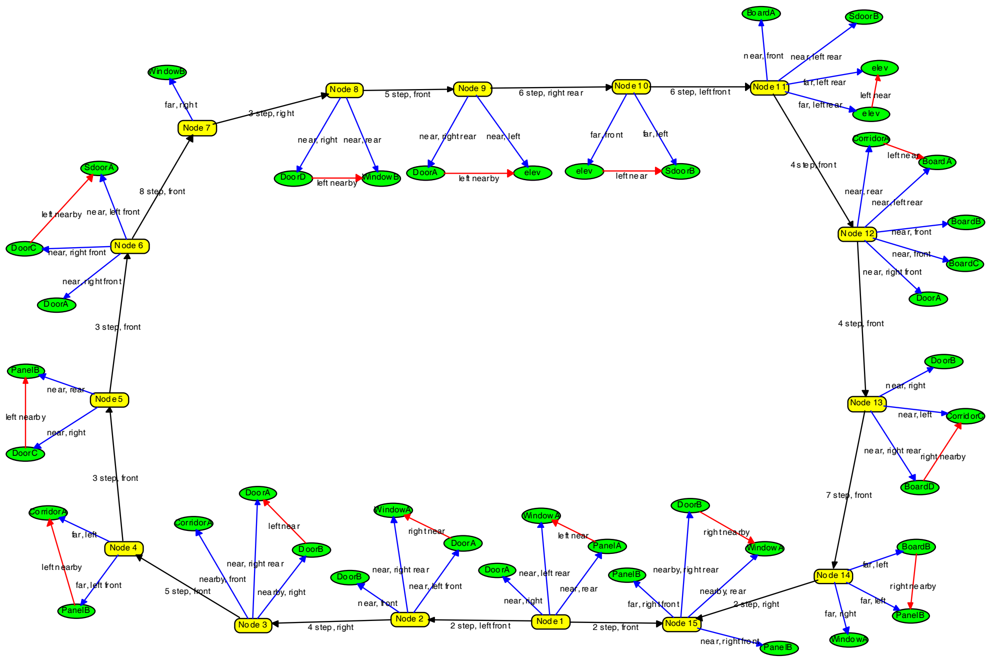


Fig. 9. Result of topological-semantic distance-map building.

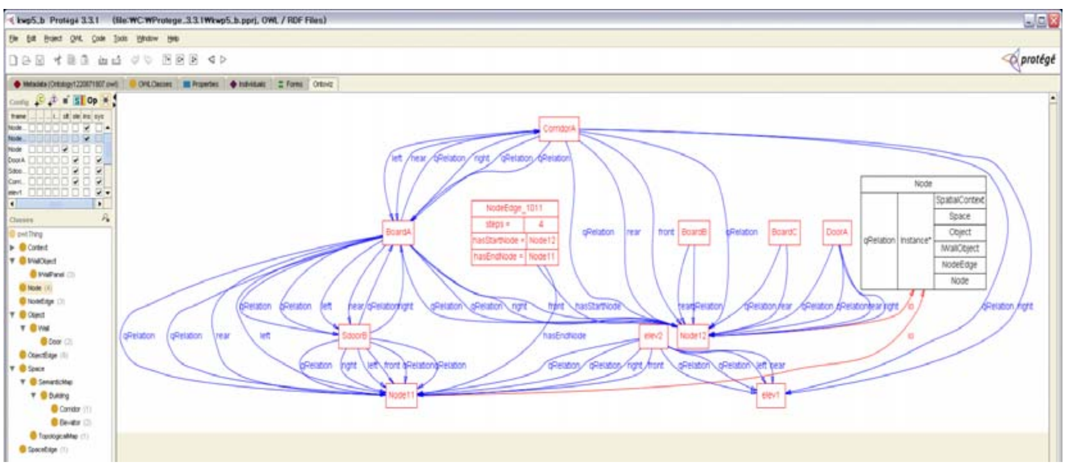


Fig. 10. Example of robot-centered ontology schema and map instance.

[7] P. Beeson, and B. Kuipers, "Towards Autonomous Topological Place Detection Using the Extended Voronoi Graph," *Proceeding of IEEE ICRA*, pp. 4384-4390, 2005.

[8] V Miro, J., Zhou, and W. Dissanayake, "Towards vision based navigation in large indoor environments," *Proceeding of IEEE the IROS*, pp. 2096-2102, 2006.

[9] S. Tully, H. Moon, D. Morales, G. Kantor, and H. Choset, "Hybrid Localization using the Hierarchical Atlas," *Proceeding of IEEE the IROS*, pp. 2857-2864, 2007.

[10] B. Krieg-Bruckner, U. Frese, K. Luttich, C. Mandel, T. Mossakowski and R. J. Ross, "Specification of an Ontology for Route Graphs," *Lecture notes in Computer Science*, pp. 390-412, 2005.

[11] J. Bateman and S. Farrar, "Modelling Models of Robot Navigation Using Formal Spatial Ontology," *Lecture notes in Computer Science*, pp. 366-389, 2004.

[12] Y. Ke and R. Sukthankar, "PCA-SIFT: A More Distinctive Representation for Local Image Descriptors," *Proceedings of Computer Vision and Pattern Recognition*, 2004.

[13] I.H. Suh, G.H. Lim, W. Hwang, H. Suh, J.H. Choi, and Y.T. Park, "Ontology-based Multi-layered Robot Knowledge Framework (OMRKF) for Robot Intelligence," *Proceeding of the IEEE the IROS*, 2007.

[14] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte Carlo Localization for Mobile Robot", *Proceeding of the ICRA*, 1999.

[15] H. Choset, K.M. Lynch, S. Hutchinson, G. Kantor, W. Burgard, L.E. Kavrakij, and S. Thrun, *Principles of Robot Motion - Theory, Algorithms, and Implementations*, MIT-Press, 2005.

[16] S. Vasudevan and R. Siegwart, "A Bayesian approach to Conceptualization and Place Classification: Incorporating Spatial Relationships (distances) to Infer Concepts," *Proceedings of the IEEE IROS, workshop From Sensors to Human Spatial Concepts (FS2HSC)*, 2007.

[17] G.H. Lim and I.H. Suh, "weighted Action-coupled Semantic Network (wASN) for Robot Intelligence," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2008.

[18] M. E. Munich, P. Pirjanian, E. D. Bernardo, L. Goncalves, N. Karlsson, and D. Lowe, "SIFT-ing Through Features with ViPR," *IEEE Robotics and Automation Magazine*, pp. 72-77, 2006.

[19] T. Bailey and H. Durrant-Whyte, "Simultaneous Localization and Mapping (SLAM): Part II," *IEEE Robotics and Automation Magazine*, pp. 108-117, 2006.

[20] N. Karlsson, E. D. Bernardo, J. Ostrowski, L. Goncalves, P. Pirjanian, and M. E. Munich, "The vSLAM Algorithm for Robust Localization and Mapping," *Proceeding of the ICRA*, 2006.