ROBUST SEMANTIC WORLD MODELING BY BETA MEASUREMENT LIKELIHOOD IN A DYNAMIC INDOOR ENVIRONMENT

Gi Hyun Lim[1], Chuho Yi[1], Il Hong Suh[2] and Seung Woo Hong[3]

[1]*Department of Electronics & Computer Engineering, Hanyang University, Seoul, Korea*

[2]*College of Information and Communications, Hanyang University, Seoul, Korea*

[3]*Department of Intelligent Robot Engineering, Hanyang University, Seoul, Korea*

Keywords: Robust knowledge instantiation, Semantic world modeling, Beta measurement likelihood.

Abstract: In this paper, a semantic world model represented by objects and their spatial relationships is considered to endow service robots. In the case of using commercially available visual recognition systems in dynamically changing environments, semantic world modeling must solve problems caused by imperfect measurements. These measurement result from variations caused by moving objects, illumination changes, and viewpoint changes. To build a robust semantic world model, the measurement likelihood method and spatial context representation are addressed to deal with the noisy sensory data, which are handled by temporal confidence reasoning of statistical observation and logical inference, respectively. In addition to the representation of a semantic world model for service robots, formal semantic networks can be exploited in representations that allow for interaction with humans and sharing and re-using of semantic knowledge. The experimental results indicate the validity of the presented novel method for robust semantic mapping in an indoor environment.

1 INTRODUCTION

Semantic world modeling is considered to provide service robots with the ability to interact with humans and share or re-use semantic knowledge (Thielscher, 2000), (Hertzberg and Saffiotti, 2008). In the real environment, a semantic world model affords to represent a dynamically changing world. Significant problems are caused by imperfect measurements, which result from variations caused by moving objects, including humans, illumination changes, and viewpoint changes (Thrun, 2002). Even if commercially available visual recognition systems are used (Munich et al., 2005), many imperfect measurements remain false positives and false negatives due to mismatches that result in false semantic world models. Relatively speaking, false positive results not a serious impediment to the visual recognition domain but can be problematic for formal logics. Insufficient facts due to false negative results can be corrected by additional true positive results, but erroneous facts due to false positive results will result in false reasoning consequences; this generates a vicious cycle, and errors are difficult to correct even with additional true negative results.

To build a robust semantic world model, the measurement likelihood function and spatial context representation are addressed to deal with the noisy sensory data, which are handled by temporal reasoning rules (Lim and Suh, 2010) using statistical observation (Park et al., 2009) and logical inference, respectively. A measurement likelihood function based on beta distribution is proposed to estimate the confidence of a sequence of sensory observations. The measurement likelihood function converts stochastically to an object recognition likelihood by matching between the model and observations. The logical modeling of temporal rules infers spatial relationships so as to check the temporal relations among observation time intervals (Allen, 1991). In addition to the representation of a semantic world model for a service robot, formal semantic networks (Suh et al., 2007), (Lim et al., 2011) can be exploited in representations that allow for interaction with humans and sharing and re-using of semantic knowledge (Yi et al., 2009b), (Yi et al., 2009a).

This paper is organized as follows: In the next section, overall architecture for semantic world modeling and localization is discussed. Formal representation of a semantic world model is described in section 3.

Lim G., Yi C., Suh I. and Woo Hong S..
ROBUST SEMANTIC WORLD MODELING BY BETA MEASUREMENT LIKELIHOOD IN A DYNAMIC INDOOR ENVIRONMENT.
DOI: 10.5220/0003690503110316
In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development* (KEOD-2011), pages 311-316
ISBN: 978-989-8425-80-5

Sections 4 and 5 explain the measurement likelihood function to determine false positive and false negative results and temporal confidence reasoning to instantiate objects and their spatial relationships, respectively. In section 6, experimental results are presented to show the validity of the presented novel method. Finally, in section 7, the conclusion and consideration are discussed.

2 OVERALL ARCHITECTURE

Figure 1 shows the overall architecture for semantic world modeling and localization. When there is an observation from the viewpoint of a service robot, its features are matched with the features of the model in object database. The likelihood of measurement is estimated using the stochastic method. By observation measurement and control, a local metric map is built from the viewpoint of the robot. A spatial context reasoner infers spatial relationships between objects from the viewpoint of the reference object, which is usually the object that if found first in the local area. The viewpoint transformation from egocentric to allocentric representation can be accomplished using logical rules for movement and rotation. When a robot moves to an-other area and finds an object, two areas are linked topologically. By using a topological-semantic distance map, global localization is made possible through the object and the spatial object contexts. In addition, the localization is processed more specifically and locally based on the observed object information around the node. Within the overall framework, the present study concentrates on robust semantic mapping, which enables humans to interact with robots.

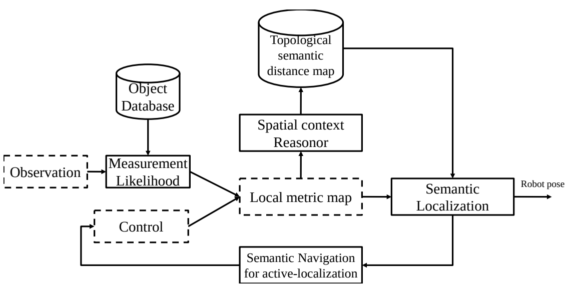

Figure 1: Overall architecture for semantic world modeling and localization.

3 REPRESENTATION OF A SEMANTIC MAP

3.1 Topological-semantic Distance Map

A topological-semantic distance map is proposed to model space by means of ontology, which ensures that only sound and complete data are asserted and propagated with ontology inference. The proposed topological-semantic distance map, which consists of spatial object contexts and spatial robot contexts, includes two types of maps: a transient local metric map and a permanent topological semantic map. A metric map is built using observation measurements. A topological-semantic map includes nodes and edges for global topological representation between nodes, objects, and their spatial relationships for local semantic representation. A node is one of the components of a global topological map that plays the role of a standard and contains information on the spatial object contexts. The 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 the qualitative bearing is the n-n bearing context. Spatial relationships are more concerned with the viewpoints of objects than the robot's own observation viewpoints. The allocentric representation is converted egocentrically in the case of localization or navigation (Yi et al., 2009b), (Yi et al., 2009a).

3.2 Spatial Object Contexts

Figure 2 shows a semantic representation consisting of observed objects and their respective spatial symbols. 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 $s = \{front, leftfront, left, leftrear, rear, rightrear, right, rightfront\}$ is the bearing of the object relative to the robot. The `o-o` relationship context denoted by $s = \{leftfar, leftnear, leftnearby, rightnearby, rightnear, rightfar\}$ is the relationship among objects.

Table 1 shows a semantic representation using symbols for all the spatial contexts in Fig. 2. Our robot localization application finds the position of the robot using only these types of semantic representations with qualitative metric data.

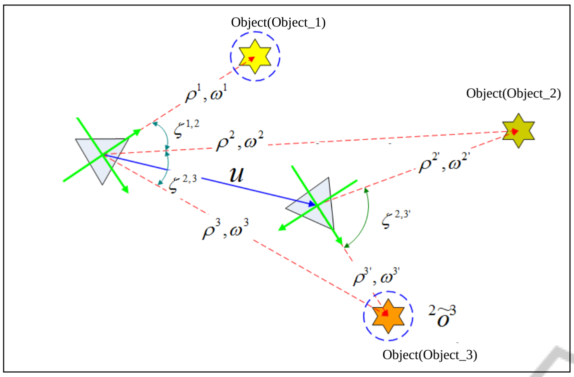

Figure 2: Spatial object contests, spatial relationships between object-based local coordinates.

Table 1: Semantic representation including all the spatial object contexts in Fig. 2.

State	Semantic representation
Previous state	nearby(o1, Robot), left front(o1, Robot), right near(o1, o2), right far(o1, o3), far(o2, Robot), front(o2, Robot), left near(o2, o1), right near(o2, o3), far(o3, Robot), right front(o3, Robot), left far(o3, o1), left near(o3, o2)
Current state	near(o2, Robot), left front(o2,Robot), right far(o2, o3), nearby(o3, Robot), right front(o3, Robot), left far(o3, o2)

4 MEASUREMENT LIKELIHOOD FUNCTION

The measurements of visual observation can be instantiated as ontology for the representation of the topological-semantic map, 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, should be filtered for robust semantic mapping. In the case of false positives, the properties are illogical, for instance, a misclassified object may make erroneous spatial relationships. Moreover, inferred erroneous facts will result in false consequences for reasoning; this generates a vicious cycle, and errors are difficult to correct, even with additional true negative results.

To address the failure of knowledge instantiation, a measurement likelihood function and a robust semantic knowledge instantiation rule is proposed to ensure the logical rigidness of robot knowledge instances.

4.1 Beta Measurement Likelihood Function

Noisy data are the result of dynamic factors and viewpoint changes. Similar to the $LeTO^2$ function (Park et al., 2009), the measurement likelihood function is introduced on the basis of beta distribution. Beta distribution can be in the form of well-applied successive independent Bernoulli trials. Each feature point of visual observation is regarded as an independent trial to determine matching. To model the measurement likelihood function, a cumulative distribution function (cdf) of beta distribution is applied. The cdf is an S-shaped function in cases where $alpha$ and $beta$ are more than 1. Given s successes in n conditionally independent trials with probability p, p should be estimated as $(s+1)/(n+2)$. This estimate may be regarded as the expected value of the posterior distribution over p, namely Beta(s + 1, n - s + 1), and then it is observed that p generated s successes in n trials.

$$f(x;\alpha,\beta) = \frac{1}{\mathrm{B}(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} = \mathrm{B}_x(\alpha,\beta), \quad (1)$$

$$F(n;\alpha,\beta) = \frac{\mathrm{B}_r(\alpha,\beta)}{\mathrm{B}(\alpha,\beta)} = \mathrm{I}_r(\alpha,\beta). \quad (2)$$

For registration, every trained object image is selected at each node by the user, and feature points of captured images are stored in the database. During matching for mapping or localization, the features extracted from an image taken at the current location of the robot are matched with those extracted from each reference image in a pre-built database.

A formal description of the beta measurement likelihood function is as follows:

$$\begin{aligned} &ML_\beta(r) = F(r;\alpha,\beta) \\ &= F(r; s+1, n-s+1) = \frac{\mathrm{B}_r(s+1,n-s+1)}{\mathrm{B}(s+1,n-s+1)} \\ &= \mathrm{I}_r(s+1, n-s+1) \end{aligned}$$

where, s = average number of matched keypoints,
n = average number of model keypoints,
r = matching ratio between model features and currently observed features,
$\alpha = s+1$,
$\beta = n-s+1$.

4.2 Likelihood Confidence Interval (LCI)

Confidence of recognition is determined by an likelihood interval-counter (γ) from the measurement likelihood for each object recognition result. An interval-counter for each object is defined on the basis of the *confidence law of inertia*, whereby a knowledge instance is assumed to persist unless there is confidence to believe otherwise. If the measurement likelihood of object `A` is x_A, then (1-x_A) is the probability that the recognition data for `A` can be false. From that, $(1\text{-}x_A)^{\gamma_A}$ can be calculated to define probability when the values of γ_A consecutive data are all false. If the result of $(1\text{-}x_A)^{\gamma_A}$ is less than 5% (0.05), then it can be said that the data have been obtained within a confidence interval (1.96σ, $P = 0.05$) of the 95% confidence level. For example, if the measurement likelihood of object `A` is 80%successively, the recognition failure rate of object `A` might be 20% (0.2). The result rate of recognition failure of two consecutive observations is 4% (0.04) and 4% is beyond the 95% confidence interval($P = 0.05$), so γ of object `A` is 2. At that time, the instance of object `A` is created and vice versa. The likelihood interval-counter using β likelihood distribution can be represented as follows:

$$\gamma_\beta = \min\{\gamma \in I| \prod_{i=1}^{n} (1 - x_{\mathrm{obj}}) \leq P\}, \qquad (3)$$

where $P = 0.05 = 1 - 95\%$ confidence level.

5 TEMPORAL CONFIDENCE REASONING

According to continuous observations from robot movement, object instances might be created or deleted whether certain number of consecutive observation likelihoods exceed the likelihood confidence interval. Time intervals of object instances which exist or not is determined by the durations between the changes of confidence. Temporal relations between intervals are inferred using temporal reasoning. The temporal relation was first proposed by Allen (Allen, 1991) and represents temporal relations using before, after, meets, met-by, overlaps, overlapped-by, and so on. Table 2 lists the rules of temporal reasoning to show the end point relations between two intervals. In the table, obj_1 and obj_2 are object instances, intervals a_l, a_m and a_n include start point a^s and end point a^e. If two intervals meet or overlap, then they are merged into one interval. The merged interval begins at the start point of the former and ends at the end point of the latter. Temporal confidence reasoning (TCR) is based on the assumption that recognized objects cannot go away and come back within a single time interval.

When an object instance of `A` is registered, if other objects are also considered to be true positive instances and to have a temporal relation of *overlapped* with object `A`, then spatial relations among the objects can be inferred. For instance, Fig. 3 presents a set of spatial relations between objects `A` and `B`. When the *is*-interval of object `B` is considered to be true, the temporal relation between a_m^+ and b_m^+ is considered to be an *overlap*. Then, the spatial relation between them can be reasoned and set using spatial reasoning. All object instances and their spatial relations can be registered in the instance database.

Table 2: Rules of Temporal Confidence Reasoning.

Temporal Relation	End Point Relations
if $obj_1 = obj_2$ **and** $a_l^{obj_1}$ **meets** $a_m^{obj_2} \Rightarrow a_n^{obj_1}$ **if** $obj_1 = obj_2$ **and** $a_l^{obj_2}$ **met_by** $a_m^{obj_1} \Rightarrow a_n^{obj_1}$	$a_l^s < a_l^e = a_m^s < a_m^e \Rightarrow a_l^s = a_n^s < a_m^e = a_n^e$
if $obj_1 = obj_2$ **and** $a_l^{obj_1}$ **overlaps** $a_m^{obj_2} \Rightarrow a_n^{obj_1}$ **if** $obj_1 = obj_2$ **and** $a_l^{obj_2}$ **overlapped_by** $a_m^{obj_1} \Rightarrow a_n^{obj_1}$	$a_l^s < a_m^s < a_l^e < a_m^e \Rightarrow a_l^s = a_n^s < a_m^e = a_n^e$

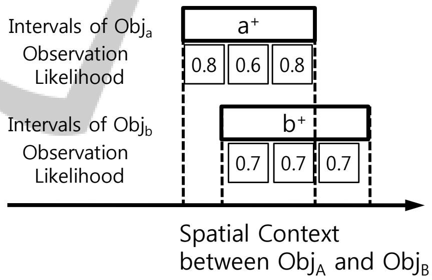

Figure 3: Temporal reasoning of spatial relations between object `A` and `B`, in which '+' denotes positive instance.

6 EXPERIMENTAL RESULTS

A Pioneer 3 AT robot carrying a single consumer-grade camera was driven around an indoor environment (14 × 10 m) to evaluate the performance of the proposed semantic mapping.

Figure 4: Examples of trained object(landmark) images.

Figure 4 shows examples of trained object images selected by the user. The camera observed 9 objects during its travel around the indoor environment. Distinctive objects such as a `printer`, a `refrigerator`, `drawers`, etc were used for object recognition.

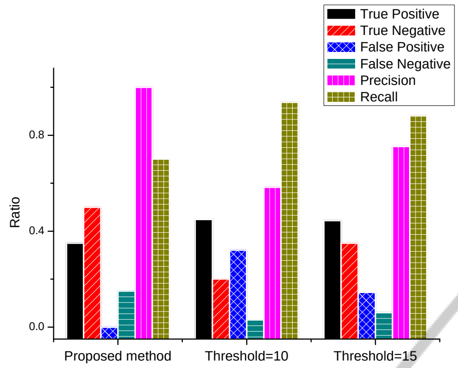

Figure 5: Experimental results.

Figure 5 summarizes the results of a three-approach, cross-validation experiment using the proposed method with a threshold of 10 and 15 correct matches. In many cases, object recognition by an Evolution Robotics Vision system was used with a threshold of 10 correct matches.

In order to confirm our results, we evaluated the performance of the system measuring its effectiveness by means of true positive, true negative, false positive, false negative, Precision, and Recall. We present the results relative to precision and false positive, which indicate overall performance. Figure 5 show that precision increases up to a level of 1. It is observed that all false positives were successfully removed from the recognition results. The experimental results reveal that the proposed method makes it possible to robustly register object instances even with an imperfect vision sensor. However, in the proposed method, false negatives cause the recall to decrease somewhat. Most of the false negatives in our method register at early instantiation. On the other hand, once a robust semantic map is built, false negatives also decrease at localization or update (Lim and Suh, 2010), (Yi et al., 2009b), (Yi et al., 2009a).

Figure 6, Figure 7, and Figure 8 show experimental results of temporal reasoning to check the validity of the relationships between intervals and statistical reasoning to determine the LCI of object recognition, where the blue line represents the trajectory which is measured using odometry.

Figure 9 illustrates a topological-semantic distance map consisting of 4 nodes (blue, rectangle) and 9 objects (pink, circle). Solid lines between nodes are the edges that represent n-n contexts of distances and bearings. The blue lines denote the r-o context and the red lines represent the o-o context.

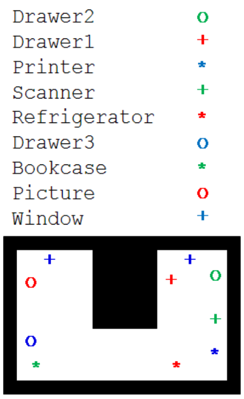

Figure 6: Experimental environment composed of nine objects.

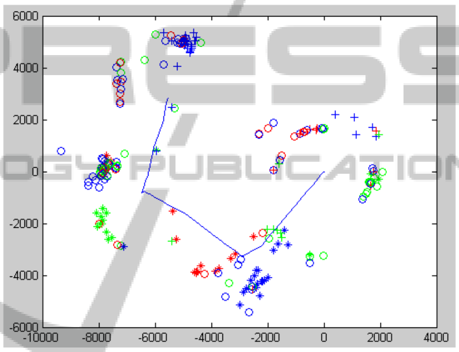

Figure 7: Specific example of mapping on imperfect sensing data.

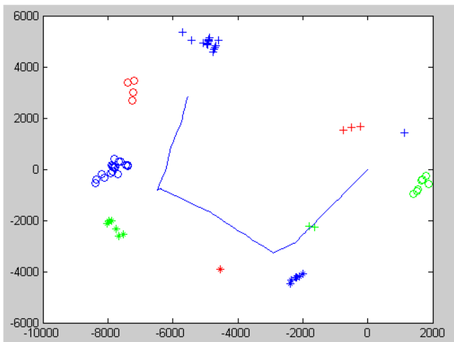

Figure 8: Result of mapping using temporal and statistical reasoning.

7 CONCLUDING REMARKS

In this paper, we proposed a robust semantic mapping method for use under conditions of imperfect object recognition. The method uses beta measurement likelihood statistical reasoning to determine the confidence interval of object recognition and tempo-

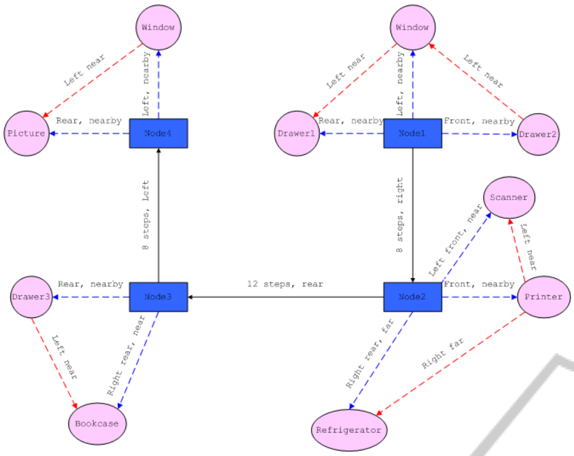

Figure 9: Result of topological-semantic distance-map building.

ral reasoning to check the validity of relationships between intervals, and represent ontological spatial relations between objects and the semantic map. Determining failures from unreliable object recognition makes it possible to dependably instantiate semantic knowledge. In our novel approach, the robot verifies the recognized objects as true or not. The experimental results indicate that all false positives in the recognition results were corrected. Therefore, a robust topological-semantic distance map, consisting of nodes, objects, and their relationshipscan be built for application in service robots.

ACKNOWLEDGEMENTS

This work was supported for the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the MKE(Korea Ministry of Knowledge Economy), and partially supported by the MKE, Korea, under the Human Resources Development Program for Convergence Robot Specialists support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2011-C7000-1001-000x).

REFERENCES

Allen, J. (1991). Planning as temporal reasoning. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, volume 13, pages 25–45. Citeseer.

Hertzberg, J. and Saffiotti, A. (2008). Editorial: Using semantic knowledge in robotics. *Robotics and Autonomous Systems*, 56(11):875–877.

Lim, G. H. and Suh, I. H. (2010). Robust robot knowledge instantiation for intelligent service robots. *Intelligent Service Robotics*, pages 115–123.

Lim, G. H., Suh, I. H., and Suh, H. (2011). Ontology-based unified robot knowledge for service robots in indoor environments. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 41(3):492 –509.

Munich, M., Ostrowski, J., and Pirjanian, P. (2005). ERSP: a software platform and architecture for the service robotics industry. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005.(IROS 2005)*, pages 460–467.

Park, Y., Suh, I., and Choi, B. (2009). Bayesian robot localization with action-associated sparse appearance-based map in a dynamic indoor environment. In *Proceedings of the 2009 IEEE/RSJ international conference on Intelligent robots and systems*, pages 3459–3466. IEEE Press.

Suh, I. H., Lim, G. H., Hwang, W., Suh, H., Choi, J. H., and Park, Y. T. (2007). Ontology-based multi-layered robot knowledge framework (OMRKF) for robot intelligence. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2007. IROS 2007*, pages 429–436.

Thielscher, M. (2000). Representing the Knowledge of a Robot. In *Principles of knowledge representation and reasoning: Proceedings of the Seventh International Conference (KR2000), Breckenridge, Colorado, April 12-15 2000*, pages 109–120. Morgan Kaufmann Pub.

Thrun, S. (2002). Probabilistic robotics. *Communications of the ACM*, 45(3):57.

Yi, C., Suh, I. H., Lim, G. H., and Choi, B. U. (2009a). Active-semantic localization with a single consumer-grade camera. In *Proceedings of the 2009 IEEE international conference on Systems, Man and Cybernetics*, pages 2161–2166. IEEE Press.

Yi, C., Suh, I. H., Lim, G. H., and Choi, B. U. (2009b). Bayesian robot localization using spatial object contexts. In *Proceedings of the 2009 IEEE/RSJ international conference on Intelligent robots and systems*, pages 3467–3473. IEEE Press.