# Ontology-based Semantic Context Modeling for Object Recognition of Intelligent Mobile Robots

Jung-Hwa Choi, Young-Tack Park
*School of Computing*
*Soongsil University*
*Seoul, KOREA*
*park@comp.ssu.ac.kr*

Il Hong Suh, Gi Hyun Lim and Sanghoon Lee
*College of Information and Communications*
*Hanyang University*
*Seoul, KOREA*
*ihsuh@hanyang.ac.kr*

***Abstract*— Object recognitions are challenging tasks, especially invisible object recognition in changing and unpredictable robot environments. We propose a novel approach employing context and ontology to improve object recognition capability of mobile robots in real-world situations. By semantic contexts we mean characteristic information abstracted from robot sensors. We propose a method to construct semantic contexts using inferences for mobile robots to recognize objects in a more efficient way. In addition, ontology has been used for better recognizing objects using knowledge represented in the ontology. OWL (Web Ontology Language) has been used for representing object ontologies and contexts. We propose a four-layered context ontology schema to represent perception, model, context, and activity for intelligent robots. And, axiomatic rules have been used for generating semantic contexts using OWL ontologies. Experiments are successfully performed for recognizing invisible objects based on our ontology-based semantic context model without contradictions in real applications.**

## I. INTRODUCTION

Object recognition in the intelligent mobile robot domain requires coping with a number of challenges, starting with difficulties intrinsic to the robot. These difficulties are caused by incomplete recognition such as imprecise perception and action, and by the limited knowledge [1], [2]. These affect the robot judgment to represent the environment and to predicate the behavior variation by user. The context modeling may lie in the imposition by the designer of an organization of the physical-world data into logical structure. We dismiss the obvious recognition of contexts in spite of hidden and partial data in the scene, preferring to promote semantic context model as the necessary basic knowledge for intelligent mobile robots.

Knowledge for robot [3] is supposed to be used to resolve the ambiguity problem when the observed positions of objects are ambiguous. Several approaches have been proposed to enhance the performance of location information.

In this paper, we represent a semantic context model on the similar principle. Context is any information that can be used to characterize the situation of an entity. An entity may be a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and application themselves [4]. In general, contexts consist of location, identity, time and activity. The context is essential for dealing with complex vision-information for intelligent robots in ubiquitous computing.

Recently, ontology has been used to represent sharable schema knowledge using classes, properties, and individuals. Ontology is useful for representing background knowledge for knowledge processing. In this paper, ontology has been used for describing various schema knowledge using taxonomy and relations between classes. Moreover, we represent all the relevant knowledge in the form of ontology for real time processing of observed objects. The idea comes from storing background knowledge in advance and using the stored knowledge for resolving ambiguities in real applications.

Contexts are represented by vocabulary defined in the ontology. Classes and properties play important roles in representing contexts. We use OWL ontology [5] to present common concepts and inference of axiomatic rules. In this way, we model the different context hierarchies more adaptively and effectively. In order to maintain coherences and consistent inferred contexts, we propose a method to automate the maintenance of visible information. Semantic context model is applicable in dynamic situation where certain sets of observed contexts are considered to be contradictory and the system must make sure that the semantic context model remains free of contradiction.

## II. RELATED WORK

Recently, many related academic research efforts and commercial reality are made for intelligent robot with vision sensor. Many knowledge-based vision systems have been designed for the application domain.

Song, Liang and Syin [6] proposed a framework for building ontology to provide semantic interpretations for image contents. Their ontology is focused on visual information such as texture, color and edge features in ontology. However the ontology is developed only for visual concepts. Wang et al. [1] proposes OWL as context ontology for modeling context and supporting logic based context reasoning. The context ontology includes person, place, entity and activity. Furthermore they developed context reasoning process. Using the mechanism, implicit context of activities such as cooking, sleeping can be deduced from explicit context such as person, device, place and so on. However they only modeled objects, places and activities, and defined rules for reasoning between them. Thus, their approach does no provide sensor-driven image or geometry information. Go and Sohn [2] studied how to model context information in a robot environment. They developed the context model using rules and ontology. The rules are used for modeling dynamic information such as user's current location, current robot dictator and so on. Their ontology is mainly used for describing static information about most parts of device, space, person and artifact. Their research focuses on modeling object, space and activities. They does not concern low level context which is generally used to recognize their context of interest in practical application.

Our goal is to acquire all the domain knowledge without requiring image processing skills. We propose a semantic context model based on ontology. The context ontology uses visual concept knowledge in order to hide the low-level vision layer complexity and to guide the rich knowledge in everyday environment. Then, we want to build knowledge bases of intelligent robot relying on the visual concept.

## III. SEMANTIC CONTEXT MODEL

A robot cannot process and understand information like humans do. The robot that acts within an environment needs machine-understandable representation of objects, in terms of their shapes, feature and usage. Our input to robots has to be very explicit so that they can handle object and determine what to do with it. We have developed a context in everyday environment using OWL ontology for everyday physical objects to support context understanding.

The need for context is even greater when we move into ubiquitous environments. The intelligent robot in ubiquitous computing have given the users the convenience that they can access whatever information and services human wants, whenever they want and wherever they are. Having the informative and accurate context, we will be able to obtain different information from the same services in different situations. Context can be used to determine what information or services to make available or to bring forward to the users.

*Context is any information that can be used to enhance robot vision related with an entity situation. An entity is nearby object, person, their relation, and changes to those objects.*

The context p($s$, $v$) is expressed using predicate logic, that is, functions over situation, that return true or false, depending on whether or not the property holds. S is subject of context; p is an attributes of s; and v is value of s. For example, the predicate left($x$, $y$) is defined to return true if object $x$ is on the left of object $y$. As this example shows, the state argument is implicit in the notation of a predicate.

A context model $CM = P(S, V)$ represents a set of contexts related to one object that invisible contexts may support. $S = \{s_1, s_2, \dots s_n\}$ is a set of context's subject; $P=\{p_1, p_2, \dots p_n\}$ is a set of attributes of $S$; and $V = \{v_1, v_2, v_n\}$ is a set of values of $S$. These model the ability to exploit resources relevant to the robot position and its ability to infer relevant information to an object. This model represents basic knowledge for robot. We work with a fixed set atoms formulas (such as encloses(*cup*, *kitchen*), These atoms stand for atomic facts which may hold of a robot, like "the cup is enclosed by the kitchen." Verbs like "encloses," "on," "left," "before," "meets," all denote predicate symbols, sometimes called propositional symbols, relationships between robots and propositions. Every variable is a term such as subject "cup" of $p$ and value "kitchen" of $p$.

There are two categories: (i) spatial context is an extension of previous research [7]. We have added context categories of presenting context to the robot. The context can be categorized into three types (with a robot, with objects and with space). An example of a context with objects might be expressed using objects relation "right", where object x is on the right of object y. (ii) temporal context is time information when defining an scene as an instance (i.e., begins, ends). For example, if we define the beginning time with the sensor information of a particular actor as the value of begins and the coordinates of the actor stay unchanged for a given time, the finishing time of the actor action is stored as the value of ending. At this time, the relations of time among actions are not considered [7]. Our context and context category for context model are in Table 1.

Semantic context is proposed to represent richer contexts using relevant ontologies. The contexts can provide richer high-level contextual information to intelligent robots. We employ context inference engine based on axiomatic method to infer ontology-based contexts based on observed context predicates. The observed context predicates need to be removed and all the inferred semantic contexts from then need to be maintained as well. For example, when an object with person enters a room, context model related to the person and object are regenerated based on ontological inferences. When the object with person leaves the room, it is necessary to retract all the semantic contexts generated. Thus, we need a system to maintain the consistencies of semantic contexts. Contexts are labeled "old_datum" or "new_datum" to indicate whether the context is observed in the situation. For instance, when a object with person enters a kitchen, the observed context node, "encloses(*object*, *kitchen*)," becomes "new_datum" in

Table 1. Context and Context Category.

| Context Type | | Predicate of Context | Description |
|---|---|---|---|
| Spatial Context | with robot | near | a robot is near a object. |
| | | far | a robot is far a object. |
| | | narrow | the road narrows for robot. |
| | | wide | it is a wide road for robot. |
| | with objects | above | one object is above another one, it is directly over it or higher than it. |
| | | below | it is inverse to "above". |
| | | over | one object is over another object. |
| | | under | it is inverse to "over". |
| | | left | one object is on the left of another object. |
| | | right | one object is on the right of another object. |
| | | beside | it is "left" or "right". |
| | | between | something is between two objects or is in between them. |
| | with place | encloses | place or object is enclosed by something. |
| | | overlap | one thing overlaps another. |
| Temporal Context | | before, after, meets, met-by, overlaps, overlapped-by, starts, started-by, during, contains, finishes, finished-by, equals | |

blackboard. Later, if the object leaves the kitchen, the observed context, "encloses(*object*, *kitchen*)" becomes "old_datum".

## IV. A Robot-centered Context Ontology And Reasoning For Semantic Context

### A. Robot-centered Context Ontology

Context predicates are generated based on vision-sensor of mobile robot. For instance, when an object movement is detected by vision sensors, a context model related to the object is regenerated. By contexts, robot refers to any information from ubiquitous sensors and infers changed information based on associated ontologies. The ontology has characteristics such as definitions of representational vocabulary, a well-defined syntax, an easily understood semantics, efficient reasoning supports, sufficient expressive capabilities that make robot recognizable.

We define major ontology as some ontology of our full robot-centered context ontology, related to vision into object ontology, a space ontology and a context ontology. Object ontology and space ontology are comprised in model ontology. Figure 1 shows an example of context ontology.
The object ontology defines every object that exists in a ubiquitous environment. For instance, it classifies objects named container (i.e., can, cup, pot, etc), furniture (i.e., table, chair, cabinet, etc), and food (i.e., rice, ice-cream, etc) in space and defines their features. A sensing data has low information (i.e., object feature) of an image including pixel information as hue, saturation, brightness, size, and shape.

An object is then inferred based on the data using axiomatic context inference engine. The inference engine analyzes the relative positions of objects based on the robot position. The information of robot is thus subsumed to the object class in order to refer data such as size and height.

The space ontology represents spatial information. As it is important to infer the movements of an actor and objects in a particular space, the space must be characterized and classified in as much detail as possible with the smallest possible zones such as path, table zone in the kitchen and corridor.

The context ontology defines various components that represent environment. It can come from the vision-data of a robot, like a human is looking around. It is the schema for context model. The ontology represents context model of objects. The context predicate is the properties for the ontology. Context ontology describes spatial context and temporal context. (i) spatial context ontology: All property values of spatial context ontology are inferred as sensing coordination of the resources in ubiquitous environment. The properties are spatial predicate of context (see Table 1).

Human and ubiquitous objects move in and out or change their position in same space. For instance, when a cup moves to another spot, context inference engine generates all the relevant contextual information by triggering context inference axioms. When it moves out the space, all of the inferred contextual information must be removed from the model. Likewise, when it moves in again, the same process will be repeated. Hence, we need to keep context dependency structure to enhance the performance of context inference engines. (ii) Temporal context ontology: We defined time ontology as OWL-time ontology [8], where the time information can be obtained regardless of user, object or location. The OWL-time ontology formalizes time by expressing it with concepts and properties. For this ontology, we reference thirteen interval relations which are proposed to formalize time by James Allen with defining relations among scenes. Temporal ontology makes it possible to obtain knowledge which is not explicitly represented in ontology by inferring relations among scenes using Allen's interval relation [9].

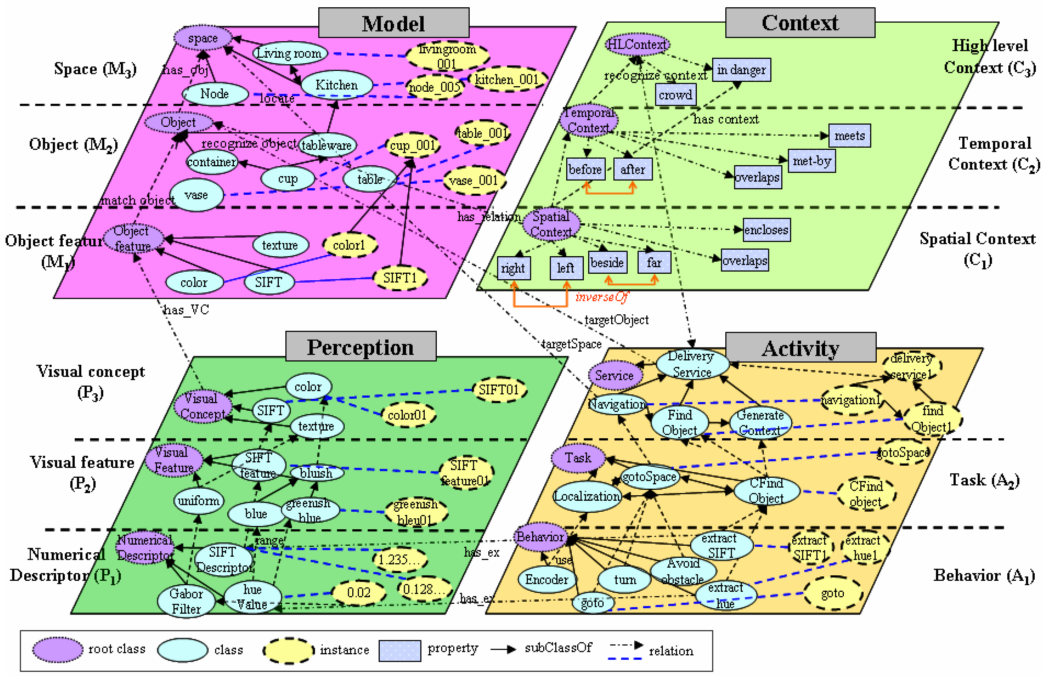


Fig. 1. An Example of Robot-centered Context Ontology Schema

### B. *Semantic Context Inference Engine Using Axiomatic Rules*

In this paper, theses robot-centered context ontology instances are used by an ontology inference engine to generate context model. The ontology inference engine reasons about the subsumption relations between contextual resources and infers new contextual information based on heuristic knowledge using axiomatic inference rules.

Our ontology inference engine is used to identify the taxonomy relations between humans and objects. This infers new knowledge using axiom provided by ontology. For instance, if we define that the domain of the property named "encloses" is "Object" class and that "Robot" class must have "encloses" property with at least one "Space" class as a range, the relation of "Robot" class's inclusion in "Object" class is defined as (Robot $\subseteq$ Object). We therefore can infer that the instance defined in "Robot" class is subsumed by “Object” class.

We classify rules into four categories, namely “create,” “retrieve,” “update” and “retract” rules, as discussed in Table 2. These APIs are implemented using a Prolog [10] with context inference engine using the rules. When retrieve-API of ontology inference engine calls, the engine

Table 2. Axiomatic Rules for Ontological Inference.

| Category | API Name and Rule |
|---|---|
| Create | createOntologyInstance(CN) → findInstance(CN, Triple) ∧ assert(new_datum(Triple)). |
| | createOntologyInstance(CN, Inst) → findInstance(CN, Inst) ∧ assert(new_datum(Inst)). |
| | setProperty(Inst, Prop, Val) → assert(new_datum(Prop(Inst, Val))). |
| Retrieve | getOntologyInstance(CN, [Inst]) → findInstance(CN, [Inst]) ∧ ssert(new_datum([Inst])). |
| | getProperty(CN, class, [Prop]) → findProperty(CN, class, [Prop]) ∧ assert(new_datum([Prop])). |
| | getProperty(Inst, instance, [Prop]) → x_position(Inst, Xpos) ∧ y_position(Inst, Ypos) ∧ theta(Inst, Theta), spatialContext(context(Inst, Xpos, Ypos, Theta, [SpatialCont])) ∧ temporalContext(context(Inst, Xpos, Ypos, Theta, [TemporalCont])) ∧ assert(new_datum([SpatialCont])) ∧ assert(new_datum([Temporalcont])). |
| | getPropertyValue(Inst, Prop, [Val]) → findPropValue(Inst, Prop, [Val]) ∧ assert(new_datum([Prop(Inst, Val)])). |
| Retract | retractOntologyInstance(Inst) → retract(new_datum(Inst)) ∧ assert(old_datum(Inst)). |
| | retractPropertyValue(Inst, Prop, Val) → retract(new_datum(Prop(Inst, Val))) ∧ assert(old_datum(Prop(Inst, Val))). |
| Update | update(Inst, Prop, Val) → retract(new_datum(Prop(Inst, [beforeVal]))) ∧ assert(old_datum(Prop(Inst, beforeVal))) ∧ assert(new_datum(Prop(Inst, Val))). |

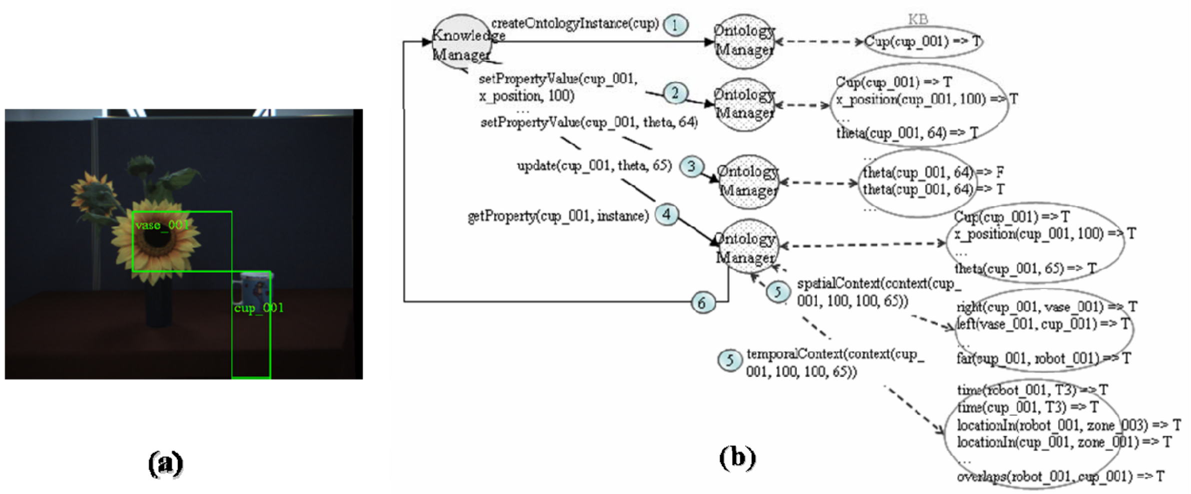


Fig. 2. Example of Semantic Context Acquisition Process; (a) Snapshot of object recognition including cup and vase (b) Semantic context acquisition process

infers semantic contexts using coordinate information of resources that are generated from ubiquitous sensors.

### C. Semantic Context Acquisition Process

As illustrated in Figure 2, the proposed semantic context acquisition process is guided by the semantic context model. It models time as a sequence of states, extending infinitely into the future. This sequence of states is simply a path of request from robot. The model allows robot to find contexts in spite of hidden and partial data. Moreover, our semantic context inference engine using axiomatic rules (see Table 2) enables robot to query through any directional reasoning between each scene.

Our inference engine is used to infer objects or some necessary information for object recognition when basic visual features (e.g., hue, saturation, brightness, brightness, size, shape, etc) are given. A simple example of Figure 2 presents how we can recognize an unknown object as a cup with objects relationship that is recognize previous scenes. In general, we wish to determine whether the object exists at the current space or not. We can use the following procedure:

1. Construct a list L of triples p(*s*, *v*), where s and *v* is any variable mentioned in the KB (working memory) and any instance in the ontology, and p is any constant mentioned in the KB and the property of the robot-centered context ontology. Our inference engine translates OWL documents into triples as facts. They are the simplest form of predicate, and held in a working memory.
2. If no such triple of visual features can be found in current space, we can use the semantic context inference engine included in ontology manager. The reasoner use collected facts in KB. In spite of unknown in current scenes, our robot is recognize hidden object via fact (of the true value) in KB.

We may assume that the query of the robot to be sent is divided according to the inference API based on ontology, which are sent sequentially. Its role is to connect the requested API, and giving them initial values of their parameters.

At a time, the cup is on the right of the vase, knowledge manager requests ontology manager to fetch ontology instance of cup and vase, and their relation (see Figure 2). The ontology manager translates ontology instances into predicate forms. When a visual recognition process is completed, the context inference engine embedded in ontology manager partially instantiates related ontology instances in real time.

In the example, the context inference engine generates the following semantic contexts which become "new_datum" in blackboard:

```
beside(cup, vase).      beside(vase, cup).
left(vase, cup).        right(cup, vase).
far(cup, robot).        above(cup, table).
encloses(cup, kitchen).
```

A few seconds later, temporal contexts occur when context inference engine generates new datum. Three temporal predicates; "time(*robot*, *T1*)," "time(*cup*, *T1*)" and "begin(*T1*)" are generated [7]. Even if some objects are not recognized due to their movement or occlusion, our system can estimate the object position by using all previous contexts telling that their *propositions* are true. The predicate "new_datum(far(*cup*, *robot*))" is generated by sensing data (see Figure 3). The followings are predicate forms remained "new_dautm" by previous contexts;

```
beside(cup, vase).      beside(vase, cup).
above(cup, table).      encloses(cup, kitchen).
```

Our inference engine can infer "behind(*cup*, *vase*)" using following a rule.

```
behind(obj1, obj2) :-
   \+ left(obj1, obj2),      \+ left(obj2, obj1),
   above(obj1, x),           above(obj2, x),
   beside(obj1, obj2).
```

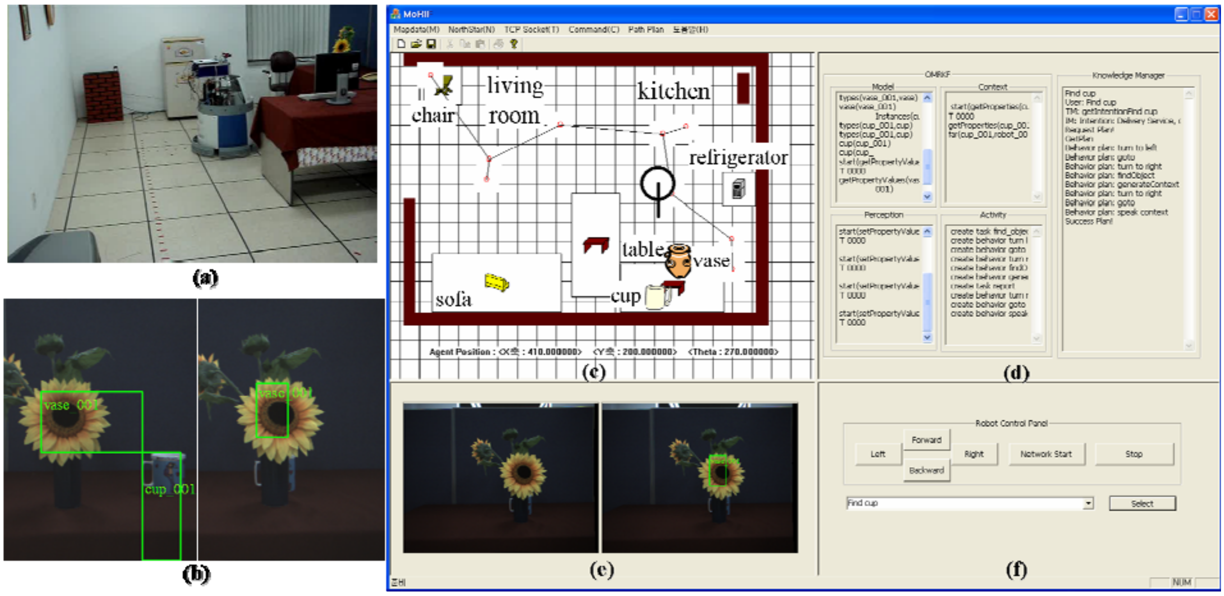


Fig. 3. Snapshots of Our Experiments; (a) 3D eye view of external camera (b) Snapshots for object recognition, in left previous case, cup and vase are recognized. But in right current case, cup is occluded by vase. (c) Top view of experiments (d) Logs and messages of robot-centered context ontology (e) Images for object recognition (f) Robot control panel and user command box

### D. *Empirical Results*

In this paper, a mobile robot is used in which a laptop and a camera are mounted. The camera gets the state of human and objects with the robot as the central position [11]. Figure 3 shows the snapshot of our experiments. Images are continuously captured by the stereo camera. The objects are firstly recognized by SIFT. And objects and their location which are occluded can be inferred by our proposed approach and compensated by our proposed axiomatic rules. Note that the proposed system can extract the relationship between objects which are not directly recognizable but inferred from relationships among the instances of spatial contexts and temporal contexts created previously (such as scene). In this experiment, the cup is on the right of the vase, while the robot is far from the object. In next scene, although the cup hides behind the vase, the robot is still able to recognize the cup because of the contexts in previous scenes.

## V. Conclusion And Future Directions

Semantic context modeling points out the necessity of situation recognition in terms of what is happening in a particular place at a particular time, or what is happening to the user. In this paper, we have presented different degrees of awareness in respect to their influences on the search space, where the context model is proven to be of tremendous importance in the recognition ability.


## Acknowledgement

This work was performed for the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by Korea Ministry of Commerce, Industry and Energy.


## References


[1] X. Wang, T. Gu, D. Zhang, H.K. Pung, "Ontology Based Context Modeling and Reasoning using OWL", *Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communication Workshops*, pp. 18-22, 2004.

[2] Y. C. Go and J. C. Sohn, "Context Modeling for Intelligent Robot Services using Rule and Ontology", *Proceedings of the 7th ICACT International Conference on Advanced Communication Technology*, vol. 2, pp. 813-816, 2005.

[3] H. J. Levesque and R. J. Brachman, "Expressiveness and Tractability in Knowledge Representation and Reasoning", *Computational Intelligence Journal*, vol. 3, pp. 78–93, 1987.

[4] A. Kumar, "Providing Architectural Support for Building Context-Aware Applications", *PhD thesis, Georgia Institute of Technology*, 2000.

[5] M. Smith, C. Welty, L. Deborah and D. McGuinness, "OWL Web Ontology Language Guide", *W3C Recommendation*, 2004.

[6] L. Song, C. Liang-Tien, and C. Syin, "Ontology for Nature-Scene Image Retrieval", CoopIS/ DOA/ ODBASE 2004, LNCS 3291, 1050-1061, 2004.

[7] J. H. Choi, I. H. Suh and Y. T. Park, "Ontology-based User Context Modeling for Ubiquitous Robots", *Proceedings of the International Conference on Ubiquitous Robots and Ambient Intelligence*, 2006.

[8] J. R. Hobbs and F. Pan, "An Ontology of Time for the Semantic Web", *ACM Transactions on Asian Language Processing (TALIP): Special issue on Temporal Information*, vol. 3, pp. 66-85, 2004.

[9] J.F. Allen, "Planning as Temporal Reasoning", *Proceeding of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, 1991.

[10] I. Bratko, "Prolog Programming for Artificial Intelligence", 3rd ed. Pearson education (2001)

[11] G. H. Lim, J. L. Chung, et al., "Ontological Representation of Vision-based 3D Spatio-Temporal Context for Mobile Robot Applications", *Proceeding of the 12th AROB Artificial Life and Robotics*, 2007.