Ontological Representation of Vision-based 3D Spatio-Temporal Context for Mobile Robot Applications

Gi Hyun Lim, Jongleul Chung,
Gwang Geun Ryu, Jong Bok Kim,
Sang Hyoung Lee, Sanghoon Lee,
and Il Hong Suh

College of Information and Communications
Hanyang University, Seoul, 133-791, Korea

Jung Hwa Choi and Young Tack Park

School of Computing
Soongsil University, Seoul, 156-743, Korea

Abstract

In this paper, we propose an ontology-base context model that consists of high level context as well as primitive spatial and temporal context. Moreover reasoning tools are used to find out not only simple contextual information such as object location, movement and distance but also hidden contextual information such as some objects disappeared by moving behind bigger objects. Also we use axiomatic rules for resolving uncertainties which might be caused by the mismatches of 3D SIFT key points. Some practical examples will be provided to show the validities of our proposed ontology-based context model.

1 Introduction

The intelligent robot needs high level perceptual tasks - context awareness[1][4][5][6], SLAM[10], object recognition. Especially contextual information is necessary for robot intelligence with which robots can recognize environments and plan their behaviors to complete missions while adapting to their environments. Such perceptual tasks are often required to be implemented by relatively inexpensive vision sensors of which visual data can be made very informative by employing many data processing algorithms. We note that vision-based context understanding system requires not only recognition of objects in the scene, but also contextual interpretation of the scene. Spatio-Temporal (ST) context is the basis for high level context understanding. However, visual data could be almost partial and occult in real environment. And such visual data processing knowledge have been specially designed for the domain-specific application. Thus, there can be hardly shared knowledge such as data structure, data processing mechanisms and rules. Using ontological representation makes it easy for intelligent robot to share its knowledge and common concepts[3][4][5][7]. Therefore, ontological representation and reasoning tools will open possibility for robots to find hidden knowledge and/or to make that knowledge is growing and reusable.

In this paper, we propose an ontology-base context model that consists of high level context as well as primitive spatial and temporal context. Moreover reasoning tools are used to find not only simple contextual information such as object location, movement and distance but also hidden contextual information such as some objects disappeared by moving behind bigger objects. Also we use axiomatic rules for resolving uncertainties which might be caused by the mismatches of 3D SIFT key points. For example, objects cannot float in the air by themselves. Instance of our proposed ontological context model will be generated based on 3D SIFT features. Specifically, after recognition of objects using SIFT features[2], primitive spatial data including location of objects, distances between objects, movement of objects are created. When the primitive spatial data are generated, an approximated center of an object is selected as a representative point for fast and efficient data processing. The compensated primitive spatial data are then clustered to instantiate primitive ST contexts. Then, higher level ST contexts are also instantiated by classifying these primitive ST contexts according to ontological taxonomy of context model. The instance is stored with a time-tag . Instances are then used to extract some missing information and/or to resolve uncertain relations by inference tools. Some practical examples will be provided to show the validities of our proposed ontology-base context model.

In the following sections, robot-centered ontology is described. Robot-centered ontology includes ontological model of perception, model, context and activity. Second, it will shown how context ontology instances are generated. Third, our ontology model will be tested on home service environment. Finally, a conclusion and consideration for future research are provided.

2 Robot-centered ontology

Robot-centered ontology is an ontological representation of robot knowledge that supports intelligent robot to perceive the environment, to model the state of the world, to plan the sequence of job, to perform the selected activity and to aware a given situation - context. Robot ontology is necessary for robot to share and reuse its knowledge, because robot perceives the environment and puts into action in a different way as human do[3]. It requires that ontological representation of robot knowledge should be suitable for its own sensors, behaviors as well as their coordination. Fig. 1 shows the architecture of ontology-based multi-layered robot knowledge framework (OMRKF), which includes 4 dimensions; KLevel, KLayer, OLayer and time and 4 levels of knowledge ($KLevel_i$); Perception(P_i), Model(M_i), Context(C_i) and activity(A_i). And those knowledge are representated by ontology.

2.1 4 Levels of Robot ontology

Robot perceives objects with its sensors, models world, plans some sequence of task, performs the task with behavior and perceive again, or robot behaves through its sensor values not with planning but with pre-programmed behaviors[11]. But, those sensor data are uncertain and partial information[10]. And service robot needs context-awareness[4], [5], [6], so that they can adapt themselves to changing situations. Context offers a few clues of the proper action selection mechanism for robot. According to the advance of context, it is necessary to develop formal context models to facilitate context representation, context sharing. Thus, OMRKF is composed of KBoards and rules, and KBoards is composed of 4 levels of knowledge perception, modeling, activity and context. each level and dimension are connected by association rules as in Fig. 1.

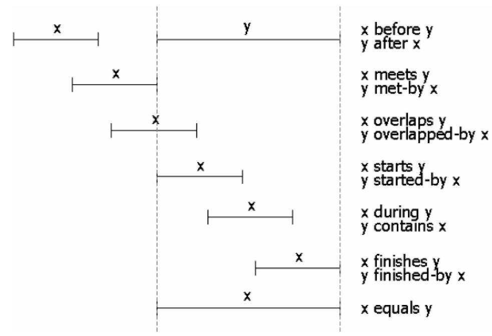

Figure 3: The Definition of Interval Relation by Allen

2.2 Context ontology

The context KLevel ($KLevel_4$) of OMRKF has 3 knowledge layers ($KLayer_{ij}$); spatial context (C_1 or $KLayer_{41}$), temporal context (C_2 or $KLayer_{42}$) and high level context (C_3 or $KLayer_{43}$) as shown in Fig.1. C_1 is is generated by primitive spatial data and includes spatial concept such as on, in, near, far, left and right. C_2 is clustered with compensated primitive spatial contest and includes spatial concepts; object-fixed, move-near, temp-moving. And C_3 is more abstracted context in specific domain with rules such as dinner, appetizer, main dish and dessert. And, basic ontological elements of OMRKF is 3 ontology layer ($OLayer_{ijk}$) such as meta ontology layer ($OLayer_{ij1}$), ontology layer ($OLayer_{ij2}$) and ontology instance layer ($OLayer_{ij3}$). Fig. 2 shows an example of OMRKF.

2.3 Temporal ontology

Time information is absolute measure which is obtained regardless of location and movement[9]. For temporal ontology, we reference thirteen interval relations as shown in Fig. 3 to define relations among actions which were proposed by J. Allen to formalize time[8]. Temporal ontology makes it possible to obtain knowledge not defined in ontology by inferring relations among locations using interval relation.

3 Instantiation of Context ontology

In this paper, vision-based objects recognition is performed by SIFT features. After recognizing objects, some instance of spatial context ontology including location of objects, distances between objects, movement of objects are created. When the spatial

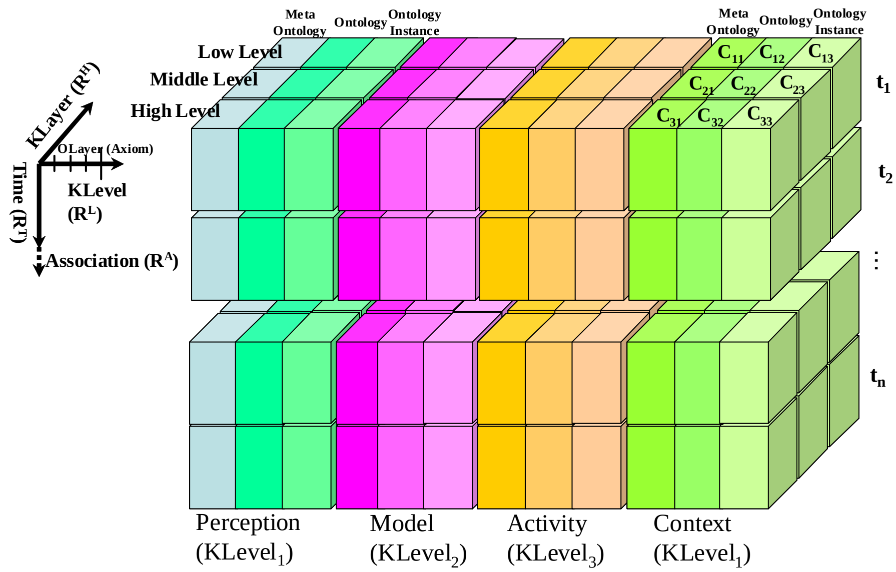

Figure 1: 4 Dimentional OMRKF

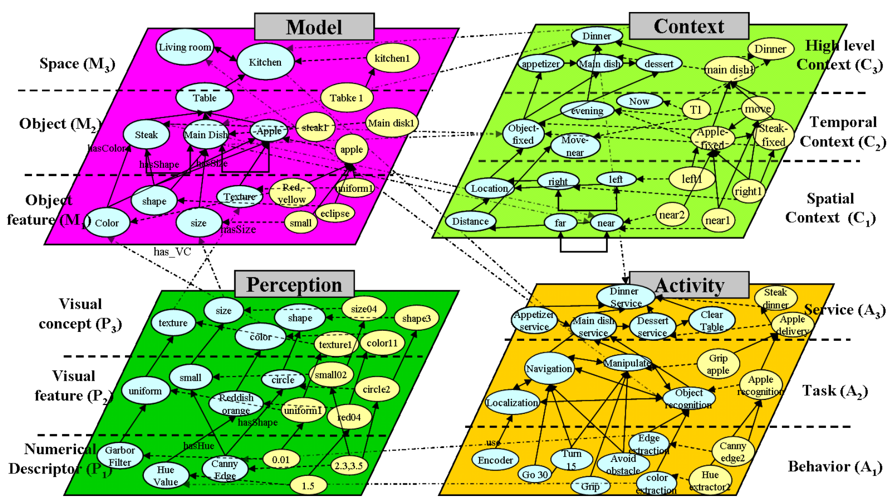

Figure 2: An Example of Robot-centered ontology Schema

contexts are instantiated ($OLayer_{413}$ or C_{13}), an approximated center of an object is selected as a representative point with axioms such as "Solid objects cannot penetrate other solid objects" and "Objects cannot float in the air by themselves". The compensated spatial context are then clustered to generate temporal context instances($OLayer_{423}$ or C_{23}). Finally, we can get high context instances($OLayer_{433}$ or C_{33}) by domain specific rules. And also we can get hidden ST contexts by inferencing pre-instantiated ST contexts.

3.1 Instantiation of Spatial Context Ontology ($OLayer_{413}$ or C_{13})

In order to recognize objects, we use a model-based approach. First of all, we use a 3D OFM (Object Feature Model) as the reference model which includes images of object, their corresponding SIFT keypoints, and representative point. By simply matching of 3D keypoints between OFM and input image, we can identify objects and obtain positions of objects with respect to robot. After identifying objects, primitive spatial data which describe geometric relations of recognized objects and are composed of location, distance, and movement data are generated. Also, the primitive spatial data are instantiated one of spatial context ontology ($OLayer_{413}$ or C_{13}).

Location Data: Location data for an object ob is defined as $R(ob) = (t, obj, R(ob))$, where t is a time when the location data is created, obj is the name of object ob, and $R(ob)$ is a position vector of object ob with respect to the robot coordinate.

Distance Data: Distance data for objects $ob1$ and $ob2$ is defined as $D(ob1, ob2) = (t, obj1, obj2, D(ob1, ob2))$, where t is a time when the distance data is created, $obj1$ and $obj2$ represent names of object $ob1$ and $ob2$, respectively. And, $D(ob1, ob2)$ is a position vector from object $ob1$ to object $ob2$ with respect to the robot coordinate.

Movement Data: Movement data for an object ob is defined as $M(ob) = (t, obj, M(ob))$, where t is a time when the movement data is created, and obj represents the name of object ob. $M(ob)$ is the vector of movement which can be simply obtained by the vector difference $R_t(ob)$ and $R_{(t-1)}(ob)$.
Table 1 shows definitions used in generation of primitive-ST data.

3.2 Instantiation of Temporal Context Ontology ($OLayer_{423}$ or C_{23})

Temporal context ontology is instantiated by inferring spatial context ontology instance such as location, distance, and movement data.

Table 1: Definition for Primitive-ST Data

Definitions	**Description**
t_s	start time
t_e	end time
$t_d = (t_e - t_s)$	time interval of interest
t_{freq}	sampling time
t_i	time acquiring i-th data
$N_{max} = \frac{t_d}{t_{freq}}$	maximum number of available data
e_R	threshold of position error
e_D	tolerance of distance error
T	trustability

Primitive location data of object $ob(pR(ob))$ shows that the object ob stays at the same location for a given period of time td. And $pR(ob)$ is defined as $pR(ob) = (id, t_s, t_e, obj, R(ob), T)$.The primitive location data can be generated as follows;

For any time $t_i \in \{0 < t_s \leq t_i \leq t_e\}$
1: If $abs(R(ob) - R_{ti}(ob)) \leq e_R$ then count = count + 1
2: T = count / Nmax
3: Generate pR(ob)

Temporal context ontology are instantiated with one or more spatial context ontology instance about the same objects. Table 2 show the spatial context and temporal context ontolgy and their FOL[7] rules.

3.3 Reasoning of context ontology

We represent OMRFK with FOL. Moreover, OMRKF includes sub-symbolic data that are seldom utilized by conventional ontology system. The data generated from robot perception or activities are numerical data, which are partial and incomplete. The probabilistic approach has dominated the solution of that case [2]. However, those systems may be application-specific, which is difficult to reuse and requires verification of the procedures. OMRKF applies sub-symbolic data to ontology-based knowledge representation, so that OMRKF can deduce hidden knowledge that is generated by a partial observation or an observation error, and make it easy to reuse and verify. Moreover, OMRKF needs rules that associate each level of knowledge, these rules enable robot to query not only unidirectional reasoning but also bidirectional reasoning. Table 2 show rules for the generation of context represented by FOL.

4 Experimental result

The ST context generation experiment was performed for objects in a refrigerator. For sequentially changing context in the refrigerator as shown in Fig.4. Mobile robot

Table 2: Rules for spatial context and temporal context ontology

Layer	Context	Logical representation
C_1	SC_1	$\forall o_1, o_2, t$ $location(o_1, t) \wedge location(o_2, t) \wedge positive((o_1.x - o_2.x), t) \Rightarrow left$
C_1	SC_3	$\forall o_1, o_2, t$ $location(o_1, t) \wedge location(o_2, t) \wedge dis\_err((o_1.x - o_2.x), t, e_D) \wedge dis\_err((o_1.y - o_2.y), t, e_D) \wedge dis\_err((o_1.z - o_2.z), t, e_D) \Rightarrow near$
C_1	SC_4	$\forall o_1, o_2, t$ $location(o_1, t) \wedge location(o_2, t) \wedge positive((o_1.y - o_2.y), t) \wedge near(o_1, o_2, t) \Rightarrow over$
C_1	SC_6	$\forall o_1, o_2, t$ $over(o_1, o_2, t) \wedge equal((o_1.y - o_2.y), (o_1.height + o_2.height)) \wedge dis\_err((o_1.y - o_2.y), t, e_D) \Rightarrow on$
C_2	TC_1	$\forall o, t, t_1$ $location(o, t) \wedge location(o, t_1) \wedge loc\_err(o, t, t_1, e_D) \Rightarrow object - fixed$
C_2	TC_2	$\forall o_1, o_2, t, t_1$ $distance(o_1, o_2, t) \wedge distance(o_1, o_2, t_1) \wedge dis\_err(o_1, o_2, t, t_1, e_D) \Rightarrow fixed - distance$
C_2	TC_3	$\forall o_1, o_2, t, t_1$ $distance(o_1, o_2, t) \wedge distance(o_1, o_2, t_1) \wedge dis\_near(o_1, o_2, t, t_1) \Rightarrow object - near$
C_2	TC_5	$\forall o_1, o_2, t, t_1$ $contain(o_1, o_2, t) \wedge contain(o_1, o_2, t_1) \Rightarrow inside$
C_2	TC_7	$\forall o_1, o_2, t, t_1, dis_1, dis_2$ $distance(o_1, o_2, t) \wedge distance(o_1, o_2, t_1) \wedge positive(dis_1, dis_2) \Rightarrow move - near$

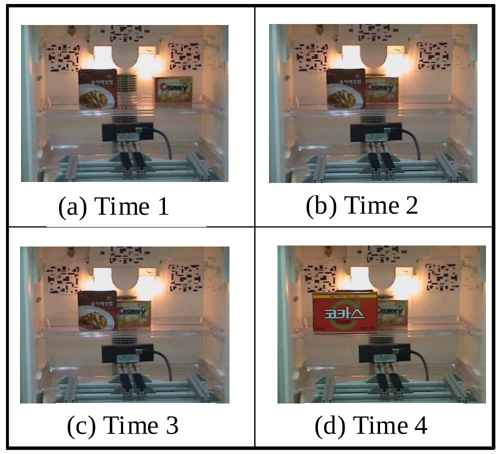

Figure 4: Sequentially Changing Contexts

may generate primitive spatial data and instantiate spatial context and temporal context ontology. At $t = t_1$, there are two objects (obj2 and obj3) apart from each other in a refrigerator(obj1). At $t = t_2$ and $t = t_3$, obj3 moves left and forward, respectively. Finally, a large object(obj4) appears in front of obj2 and occludes it.

Fig.5 shows the display of our context generation system. In the left upper corner, images captured by stereo camera attached on top of the mobile robot are displayed continuously. In the right upper corner, the location of objects, where the location is obtained by our proposed point-based approach and compensated by our proposed axiomatic rules, can be also observed. In the lower part of Fig.5, ST Context predicates generated at experiment are displayed. And ST Context Predicates for situations inside refrigerator are listed in Table 3. Note that the proposed system can extract the relationship between objects which are not directly visible by inferencing relationship among the instance of spatial context and temporal contexts created at previous times. In this experiment, although obj2 is occluded by obj4, contexts such as $front(obj4, obj2)$ and $front(obj2, obj3)$ were generated.

5 Conclusion

We proposed ontology-based context model for the household service robot, and the ontology is represented by FOL. This model allow robot to find contexts in spite of hidden and partial data. Moreover OMRKF enable robot to query through any directional reasoning between each layer as well as between each level of knowledge with small number of clues.

For a future work, we would like to extend our context model to include knowledge to handle uncertain and partial data. We also would like to represent our knowledge model as OWL/SWRL which is known to be decidable.

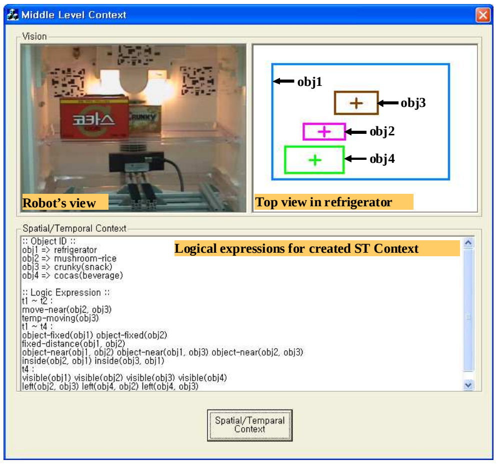

Figure 5: Shapshot of Experiments: Inside a Refrigerator

Table 3: ST Context Predicates generated at Experiment

Time	**ST Context Predicates for situations inside refrigerator**
$t_1 \sim t_2$	move-near(obj2, obj3), temp-moving(obj3)
$t_1 \sim t_4$	object-fixed(obj1), object-fixed(obj2), fixed-distance(obj1, obj2), object-near(obj1, obj2), object-near(obj1, obj3), object-near(obj2, obj3), inside(obj2, obj1), inside(obj3, obj1)
t_4	visible(obj1), visible(obj2), visible(obj3), visible(obj4), left(obj2, obj3), left(obj4, obj2), left(obj4, obj3), front(obj2, obj3), front(obj4, obj2), front(obj4, obj3), near(obj1), near(obj2), near(obj3), near(obj4)

5.1 Acknowledgements

This work was performed by a grant of the Second Stage of BK21 Research Division for Advanced IT Education Program on Industrial Demand. and 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] J. Chung, H.W. Kim, S. Lee et al, "Vision-based 3D Spatio-Temporal Context Generation for Mobile Robot Applications," The 21st International Technical Conference on Circuits/Systems, Computers and Communications(ITC-CSCC2006), pp I-77 I80, 2006

[2] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, Vol.60, pp.91-110, 2004.

[3] E. Bozsak, M. Ehrig, S. Handschuh et al, "KAON ? Towards a Large Scale Semantic Web," *Lecture notes in computer science*, No. 2455, 2002, pp. 304-313.

[4] W. Hwang, J. Park, H. Suh et al. "Ontology-based Framework of Robot Context Modeling and Reasoning for Object Recognition," to appear in /em Lecture Notes in Computer Science, 2006.

[5] X.H. Wang, D.Q Zhang, T. Gu and H.K. Pung, "Ontology Based Context Modeling and Reasoning using OWL," Proc. of the Second IEEE Annual Conference on Pervasive Computing and Communication Workshops (PERCOMW04), 2004.

[6] K. Henricksen, J. Indulska, and A. Rakotonirainy, "Modeling Context Information in Pervasive Computing Systmems," *Lecture notes in computer science*, 2002, no. 241, pp. 167-180.

[7] I. Bratko, *Prolog programming for artificial intelligence*, 3rd ed. Pearson education, 2001, pp. 57.

[8] J.F. Allen, and H.A. Kautz, *A model of naive temporal reasoning*, In Formal Theories of the Commonsense World. J. R. Hobbs and R. C. Moore, eds., Ablex., pp.251-268, 1985.

[9] J.R. Hobbs, J. Pustejovsky, "Annotating and reasoning about time and events," In Proceedings of AAAI Spring Symposium on Logical Formalizations of Commonsense Reasoning, 2003.

[10] S. Thrun, W. Burgard and D. Fox, *Probabilistic Robotics*, MIT Press 2005, ch. 1.

[11] R.C Arkin, *Behavior-based Robotics*, MIT Press, 1998.