Object Co-occurrence Graph for Object Search in 3D Environment

Puwanan Chumtong, Yasushi Mae, Kenichi Ohara, Tomohito Takubo, Tatsuo Arai

Graduate School of Engineering Science, Osaka University
1-3, Machikaneyama-cho, Toyonaka, Osaka, Japan 560-8531
(Tel: 81-66-850-6365)
{c_puwanan, k-oohara, takubo}@arai-lab.sys.es.osaka-u.ac.jp
{mae, aria}@sys.es.osaka-u.ac.jp

Abstract: We propose a method of using object co-occurrence graph for mobile service robots to search for small-scale objects in 3D environments without pre-defined map. Object co-occurrence graph describes co-occurrence relations between objects in scenes. The presence of other large objects in the environment, which are easier to be found, provides cues to search for a small object related to the large object. If the target object is often near to a large object, robot firstly finds the large object, and then searches around the large object for the target object. Thus, using object co-occurrence graph makes object searching task easier and more efficient. The object co-occurrence graph is automatically constructed by many tagged images from the WWW. Our experiment shows a robot searching for a target object using object co-occurrence graph.

Keywords: Object Co-occurrence Graph, Object searching, Mobile robot.

I. INTRODUCTION

The object search is an essential task for mobile robot which helps human in everyday environment. Many methods for object search in 3D environment have been proposed. For example, David et al. [1], Masuzawa et al. [2] and Jeremy et al. [3] use color information to verify the candidate of target object in the scene. However robot may face some difficulties to find the small target object in the case of occlusion problem and insufficient current image's resolution.

On the contrary, some methods focus on existing cues in the observing environment for object search. Dominik et al. [4] proposed the usage of object's relationship to search for target objects in unknown environment. Cipriano et al. [5] use the position probability of the observing map and found objects in their searching task. Ksenia et al. [6] utilize position probabilistic with the given hint in motion planning for searching task. Importantly, the utilizing of cues can decrease the difficulty of searching task when target object, especially small object, cannot be observed in the current robot's view.

In our approach, we focus on employing object's co-occurring relation as our cues because this relation is represented for the frequency for how often object A and object B co-occurr together. Thus we propose the Object Co-occurrence Graph (OCG) describing the co-occurring property between objects. That is, we utilize the existing of cue objects to search for target object in unknown environment. This idea can be illustrated in Fig. 1. The benefit of using object co-occurring cues among previous mentioned researches is that our approach can be easily constructed and applied as robot's knowledge.

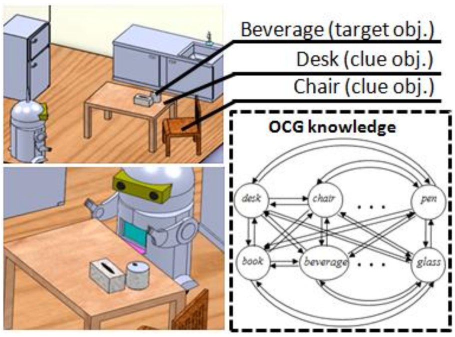

Fig. 1: The utilizing of cues in observing environment for robot's searching task

II. OBJECT CO-OCCURRENCE GRAPH

In our approach, the concept of Object Co-occurrence Graph for robot's searching task can be illustrated in Fig. 2. One arbitrary object o_i (e.g., desk, chair, etc) may have many related objects according to its reference database. Moreover, in term of co-occurring relation, each related objects of o_i may have some relationship with any objects in OCG database. For each couple of objects o_i and o_k, we will obtain 2 values; $p(o_i \mid o_k)$ and $p(o_k \mid o_i)$. At here, $p(o_i \mid o_k)$ is the probability to find target object o_i when cue object o_k is found and vice versa.

In order to utilize these probability values, we select the next visiting object candidate based on the priority ranking among others in database. That is, if our target object, o_t, is not observed while cue objects are found, robot then can determine next visiting object o_j by using the following concept:

$$\hat{o}_j = \underset{o_j \in Obj}{argmax}\, p(o_t \mid o_j) \tag{1}$$

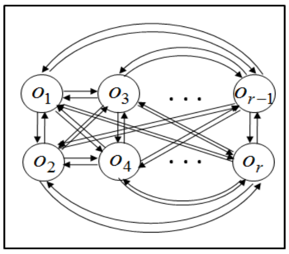

Fig. 2: The main concept of OCG structure

We know that $Obj = \{o_1, \ldots, o_r\}$ where r is number of objects. It is assumed that the higher possibility value, the higher chance to find target object nearby that cue.

III. INTERNET-BASED OCG CONSTRUCTION

The OCG can be constructed by manually learning from observing actual situation. However this learning method is not appropriate since it is time consuming. As a learning method, one interesting research [7] uses the object co-occurrence gathering from many internet images as prior knowledge for robot. As a result of this, the expected length of the path to the object is minimized. Due to benefits of internet data, the similar concept was also applied in our work. The different point of our approach from others is that our approach does not require pre-defined map for robot's searching.

To construct OCG, we employ image hosting website that contains a form of metadata as our reference database. Instead of employing well-known web search engine such as Google, the related work [8] describes the difficulty of using images from Google search as training data because most of returned images may be visually unrelated to the intended keyword. Moreover Google's images do not contain tagged information. Thus we use website such as Flickr that gives us word-tagging images provided by millions of internet users. Most of those provided images are related to searching keyword since its image searching process is based on word-tagging of each image.

The concept of OCG construction is shown in Fig. 3. The name of o_i is used as keyword to search for relevant images within main page, $W(o_i)$. Based on tags of these images in sub-webpage, $w_{k_i,i} \in W(o_i)$, the standalone object occurrence and object co-occurrence will be obtained. In this work, we have proposed 2 methods; dependent and independent co-occurring method, for OCG construction. For dependent co-occurring method, possibility values depend on tagged information in all searching images but the adding of new item or the deleting of existing item can affect the changing of possibility values of all object couple in database. Thus the recalculation of all values is needed. On the other hand, for independent co-occurring method, it is assumed that probability values between couples are independent. Thus probability values of new adding items with others can be separately calculated. The

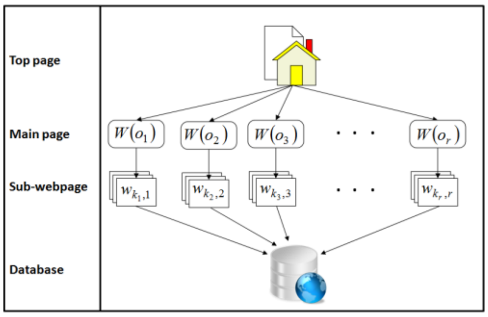

Fig. 3: Proposed method for OCG construction

computation of these 2 construction methods are based on different searching concept as follow.

1. Dependent Co-occurring based Construction

This method constructs database based on the relationship of all objects. By using name of o_i as searching keyword, we will obtain $w_{k_i,i} \in W(o_i)$. The standalone occurrence $n(o_i)$ and co-occurrence $n(o_i \cap o_j)$ are counted from all images, $\forall w_{k_t,t} \in W(o_t) : t = \{1, \ldots, r\}$. In order to gain the co-occurrence possibility value, we employ the definition of conditional probability:

$$p(o_i \mid o_j) = \frac{p(o_i \cap o_j)}{p(o_j)} \tag{2}$$

We employ the following definition:

$${}_nC_k = \binom{n}{k} \tag{3}$$

Thus, we will have ${}_rC_2$ pairs of object co-occurring couple. We will finally have $2 \cdot ({}_rC_2)$ possib ility values for OCG database.

2. Independent Co-occurring based Construction

In the case of considering relationship of $\rho - th$ pair between o_i and o_j, the standalone occurrence $n_\rho(o_i)$ and $n_\rho(o_j)$ are calculated from images $w_{k_i,i} \in W(o_i)$ and $w_{k_j,j} \in W(o_j)$. At here, we define that $\rho \in \{1, \ldots, {}_rC_2\}$ because we will obtain ${}_rC_2$ pairs of object co-occurring couple. We also can obtain $n_\rho(o_i \cap o_j)$ from those images. When the relationship between o_i and o_c is needed, we just focus on the information in $w_{k_i,i}$ and $w_{k_c,c}$. Unlike previous mentioned method, recalculation of entire values is not necessary.

IV. TRAINING RESULTS AND DISCUSSION

We design the training set based on images in Flickr website. The word "*period*" defines the image-searching of $\forall o_i \in Obj : i = \{1, \ldots, r\}$ within M images. We define that:

$$M = \sum_{i=1}^{r} m_i \tag{4}$$

At here, m_i can be varied according to the number of returned images in 1 webpage when name of o_i is used as searching keyword. Thus the searching of M

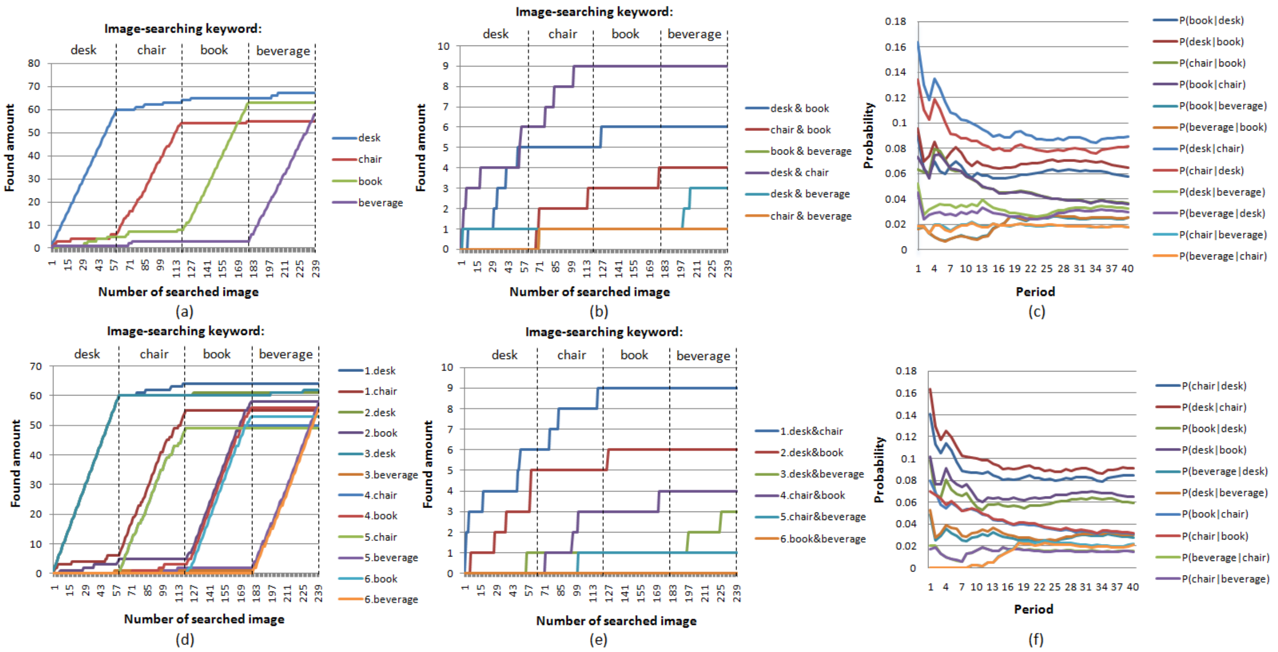

Fig. 4: Training results for OCG construction. Training results of dependent co-occurring method are shown in (a), (b), and (c) while training results of independent co-occurring method are shown in (d), (e), and (f).

images is so called one period. In our experiment, we take a consideration to 4 sample objects which are *desk*, *chair*, *book*, and *beverage*; that is, we will have $r = 4$.

For our OCG construction, we train data up to 40 periods and examine how possibility values change as the number of searching images increases. As can be seen in Fig. 4, we also provide the following definition:

- 1st phase: searching keyword of *desk* (o_1)
- 2nd phase: searching keyword of *chair* (o_2)
- 3rd phase: searching keyword of *book* (o_3)
- 4th phase: searching keyword of *beverage* (o_4)

1. Experimental Results of Dependent Co-occurring-based Construction

According to the method in section III-1, During 1st phase of Fig. 4(a), it can be noticed that $n(o_1)$ rapidly increases until the number of searching images reaches m_1 images because the name of o_1 is used. During 2nd phase, $n(o_1)$ slowly increases although name of o_2 is used. This shows that there are some co-occurrences of o_1 and o_2 during 2nd phase. The co-occurrence of *desk* (o_1) & *chair* (o_2) can be easily observed in Fig. 4(b). On the contrary, $n(o_2)$ does not change during the 4th phase of Fig. 4(a). This shows that there is no co-occurrence of o_2 and o_4 during this phase. We can also see the co-occurrence of *chair* (o_2) & *beverage* (o_4) in Fig. 4(b) whose data-line remains constant during 4th phase.

For 40 period-results in Fig. 4(c), data-lines of probability values, $p(o_i \mid o_j)$ where $o_i, o_j \in Obj : i \neq j$, provide less fluctuation as the number of searching period increases. In our experiment, the data training of 40 searching periods returns us 9,571 related images. It can also be noticed that data-lines provide less fluctuation when number of searching images is more than 4560 images (after 19th period).

2. Experimental Results of Independent Co-occurring-based Construction

According to the method in section III-2, we obtain 6 pairs of co-occurring relationship as can be seen in Fig. 4(d). Their change is based on the appearance of o_i in related images. For the 1st pair ($\rho = 1$) whose members are *desk* and *chair*, $n_1(o_1)$ changes during the 1st and 2nd phase where we use searching keywords of *desk* and *chair*, respectively. From Fig. 4(e), $n_1(o_1 \cap o_2)$ increases during 1st and 2nd phase and remains constant in other phases because this method concerns only images of $w_{k_i,i} \in W(o_1)$ and $w_{k_j,j} \in W(o_2)$.

According to Fig. 4(f), the training process of 40 periods returns us 9,578 searching images. It can be noticed that data-lines in Fig. 4(f) have similar characteristic as data-lines in Fig. 4(c).

3. Comparison Between 2 Construction Methods

The probability results of each couple of objects after 40-period training are shown in Table 1. Although the numerical results of probability values between 2 proposed methods are different, their ranking order of hypothesis cue objects convey the same ranking order and they indeed accord to the fact of object's relationship in realistic environment. From Table 1, it can be noticed that this knowledge lets robot know visiting location for object search. That is, if robot is assigned to search for *book* but it cannot be found in the

Target Object	Cue Object	Dependent Co-occurring	Independent Co-occurring
desk	*chair*	0.0890	0.0910
	book	0.0643	0.0649
	beverage	0.0325	0.0298
chair	*desk*	0.0815	0.0846
	book	0.0364	0.0320
	beverage	0.0176	0.0149
book	*desk*	0.0575	0.0598
	chair	0.0355	0.0309
	beverage	0.0250	0.0208
beverage	*desk*	0.0296	0.0279
	book	0.0255	0.0219
	chair	0.0175	0.0151

Table 1: The comparison of experimental results between 2 proposed methods

1st scene, robot then has to determine where it has to go next to search for it. In case that *desk* and *chair* are observed in the 1st scene, robot will know that *desk* is 1st place to be visited while *chair* is the 2nd place to be visited. This robot's knowledge is obtained from OCG.

Although both construction methods provide us the same meaningful results, the usage of independent will be more appropriate for real experiment because it can be easily adjusted when new object is added or existing object is removed from current database.

V. OBJECT SEARCH USING OCG

We show a simulation of object search using OCG and example of object recognition. In our object recognition method, range data and color image gathering from single stereo camera are employed as our input data for robot's exploration. The experimental set up can be viewed in Fig. 5(a). In this test, *book* is assigned as target object for robot's searching task. The robot employs well-known SIFT [9] for target object's detector. As can be seen in Fig. 5(b), robot cannot observe *book* in the 1st scene due to low image's resolution and occlusion from other objects. However robot found some cues; *desk* and *chair*, labeling with red ellipse and green ellipse, respectively. Cue objects, considered as large objects, are detected based on range data and color segmentation result [10]. Consequently robot obtains cue objects in the observing environment. Thus, robot can determine the sub-goal's location based on OCG knowledge. As a result of this, robot chooses to visit the desk's location and continue searching for *book* as can be seen in Fig. 5(c). From Fig. 5(d), it can be noticed that robot can finally verify the existing of *book* which is placed on *desk* (cue object).

VI. CONCLUSION

This paper has described the usage of OCG as robot's knowledge for searching task in unknown environment. The OCG is automatically constructed based on searching images within Flickr website. Two proposed methods for OCG construction; dependent and independent co-occurring methods are also presented. In this paper, we also show the successive result of robot's simulation of object search using OCG without pre-defined map.

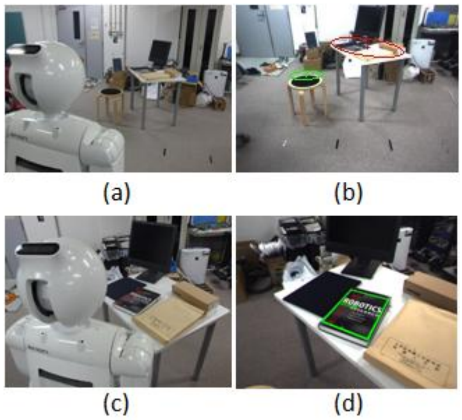

Fig. 5: Simulation of object search using OCG. (a) shows initial robot's position. (c) shows robot's positions after approaching to *desk*. (b) and (d) show robot's views with recognition results.

REFERENCES

[1] David Meger, Per-Erik Forss´en, Kevin Lai, Scott Helmer, Sancho McCann, Tristram Southey, Matthew Baumann, James J. Little, David G. Lowe, and Bruce Dow, "Curious George: An Attentive Semantic Robot", in Robotics and Autonomous Systems, 2008

[2] Hiroaki Masuzawa amd Jun Miura, "Observation Planning for Efficient Environment Information Summarization", in International Conference on Intelligent Robots and Systems (IROS), Louis, USA, 2009

[3] Jeremy Ma and Joel W. Burdick, "A Probabilistic Framework for Stereo-Vision Based 3D Object Search with 6D Pose Estimation", in IEEE Internal Conference on Robotics and Automation (ICRA), Alaska, USA, 2010

[4] Dominik Joho and Wolfram Burgard, "Searching for Objects: Combining Multiple Cues to Object Locations Using a Maximum Entropy Model", in IEEE International Conference on Robotics and Automation (ICRA), Alaska, USA, 2010

[5] Cipriano Galindo, Juan-Antonio Fernández-Madrigal, Javier González, and Alessandro Saffiotti, "Robot task planning using semantic maps", in Robotics and Autonomous Systems, Volume 56, 2008, pp.955-966

[6] Ksenia Shubina and John K.Tsotsos, "Visual search for an object in a 3D environment using a mobile robot", in Computer Vision and Image Understanding, Volume 114, 2010, pp.535-547

[7] T. Kollar and N. Roy, "Utilize object-object and object-scene context when planning to find things", in IEEE Internal Conference on Robotics and Automation (ICRA), Kobe, Japan, 2009

[8] R. Fergus, L.Fei-Fei, P. Perona, A.Zisserman, "Learning Object Categories from Google's Image Search", in International Conference on Computer Vision (ICCV), Beijing, Chaina, 2005

[9] David G.Lowe, "Local Feature View Clustering for 3D Object Recognition", in Computer Vision and Pattern Recognition (CVPR), Hawaii, USA, 20012-15

[10] Pedro F. Felzenszwalb and Daniel P. Huttenlocher, "Efficient Graph-Based Image Segmentation", in International Journal of Computer Vision, Volume 59, 2004, pp.167-181