An active reconstruction algorithm based on partial prior information

Yanzi Kong1,2,3,4,5, Feng Zhu1,2,4,5, Yingming Hao1,2,4,5, Haibo Sun1,6, Yilin Xie1,2,3,4,5 and Zhiyuan Lin1,2,3,4,5

Abstract
Active reconstruction is an intelligent perception method that achieves object modeling with few views and short motion paths by systematically adjusting the parameters of the camera while ensuring model integrity. Part of the object information is always known for vision tasks in real scenes, and it provides some guidance for the view planning. A two-step active reconstruction algorithm based on partial prior information is presented, which includes rough shape estimation phase and complete object reconstruction phase, and both of them introduce the concept of active vision. An information expression method is proposed that can be used to manually initialize the repository according to specific visual tasks, and then the prior information and detected information are used to plan the next best view online until the object reconstruction is completed. The method is evaluated with simulated experiments and the result is compared with other methods.

Keywords
Active reconstruction, information gain, partial prior knowledge, rough shape estimation, view planning

Date received: 17 October 2019; accepted: 8 January 2020

Introduction
Geometric modeling of objects is one of the important research areas in the field of robot vision. It refers to the technology that obtains the 3-D geometric information of objects through visual methods. Object modeling is widely used in the fields of object inspection, target tracking1,2 or to generate area maps for environmental modeling,3 and so on. The data obtained from one view are often incomplete for object reconstruction, and large-area defects may even occur. Therefore, it is necessary to obtain and fuse the information from different views to achieve the complete geometric reconstruction. In traditional target modeling, robot motion is passive, and the motion path is set artificially in advance by humans or controlled in real time.4 Then, a multi-viewpoint modeling is implemented using a motion mechanism with a specific motion function.5,6 To accomplish the vision task with fewer views and shorter motion paths, the concept of active object reconstruction is introduced. It means a reconstruction process in which the next best view (NBV) is planned actively to achieve the above optimization targets.

1 Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China
2 Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang, China
3 University of Chinese Academy of Sciences, Beijing, China
4 Key Laboratory of Opto-Electronic Information Process, Shenyang, China
5 The Key Laboratory of Image Understanding and Computer Vision, Shenyang, China
6 Northeastern University, Shenyang, China

Corresponding author:
Feng Zhu, Shenyang Institute of Automation, Chinese Academy of Sciences. No.135 Chuangxin Road, Hunnan District, Shenyang, 110179 Liaoning, China.
Emails: fzhu@sia.cn; 1320574041@qq.com
The model space is initialized at the beginning of the active object reconstruction, and then the following loop is executed before the termination condition is reached: (1) moving sensor to a specified location to acquire the object information, (2) updating the model space by merging the current perception information with the existing information, and (3) reasoning about the unobserved region and planning the NBV with the reconstructed model. Object reconstruction and sensor moving parts are not taken into account in active object reconstruction. We only focus on the model prediction and NBV determination. The unknown area of the object is generally derived from the continuity of the surface. The NBV can be obtained through formula analysis method or generation test method. The first one means that the view is directly calculated according to object’s geometric features, and the second is to generate a series of candidate views, and the view whose optimization function can get the highest score is selected as the NBV. To speed up the object reconstruction process, we need to make the next view contribute more to the reconstruction. It means that more unknown information should be received at the NBV, and it can be measured by the information gain (IG) of the view.

Prior information facilitates the completion of various visual tasks, and object’s partial priori information can provide some guidance for real-time view planning, although it is not sufficient to preplan all views and the motion path of the sensor before reconstruction. Our work aims to reconstruct the geometry with relatively regular shape. We know the general type of the target but do not get the knowledge of its specific characteristics such as the size before modeling. Since a kind of geometry has common geometric features, we can fill these features into a prior information library in a certain form. The movement of the view can be guided according to the prior information to estimate the rough shape of the target in the modeling process. After obtaining the rough shape of the object, candidate views’ IG can be calculated using the predicted volumetric model. View planning in our system is embodied in two stages. The first phase is for rough geometric shape estimating, and view planning at this stage is based on the prior information library and the inference engine, the other is the complete reconstruction phase, and view planning at this stage is based on the IG calculation of views.

Contributions

In this article, an active object reconstruction method that does not have much restricts on objects’ geometry size is proposed. It guides the sensor to autonomously detect the rough geometric features of the target through reasoning the prior information library. Then it completes the object reconstruction by the IG calculated method based on the rough shape of object. The main contributions of this article are listed as follows:

1. A novel active reconstruction framework for relative regular geometry is proposed.
2. A scalable prior knowledge base is created and an inference algorithm for rough shape estimation is proposed based on it.
3. An IG formulations based on prediction surface is proposed.

Organization

After the method overview in the third section, we introduce the establishment of the prior information base and the reasoning of the object’s rough geometric shape in the fourth section. The proposed IG formulation is presented in the fifth section. In the sixth section, we show the experimental results and comparison to other works, and a conclusion is proposed in the final section.

Related works

Connolly et al. first proposed the concept of view planning in 1985, and articles in this field have been published over the years since Aloimonos et al. published an article about active vision in 1988. Among them, a review article on active object reconstruction published by Scott et al. in 2003, and another one on active vision published by Chen et al. in 2011 summarized the latest research results. Based on the type of modeling task, active object reconstruction can be divided into three categories: model-based, non-model-based, and semi-model-based methods which differ in the amount of prior information about the object.

Model-based methods consider that geometric features of the target are almost completely known. A series of views can be preplanned based on the prior information, which is independent of the real-world detection. Non-model-based methods have no prior information before the modeling task, and the view planning can only rely on the information detected by sensors in real time. Non-model-based methods are widely used and more complex in practical scenarios, so they get more attention than other methods. When we take consideration of the model type and the information used in view planning, the non-model-based methods include three types: surface-based methods, search-based methods, and combined methods.

The object is represented with a model such as point cloud or triangular mesh in surface-based methods, which extract the surface information of a local model that has been obtained so far to estimate the shape of the unknown region for the NBV determination. Chen and Li used the evaluation function to sort the edge points and selected the smoothest region as the detection region of the next view. Li and Liu sliced the object into a series of closed B-spline curves, and the view that has the largest information entropy is selected as the NBV. A volumetric model is used in search-based methods to select the NBV by
defining appropriate optimization functions based on the occupancy probability of each voxel. Isler et al.\textsuperscript{17} and Vasquez-Gomez et al.\textsuperscript{18} proposed various ways to express the voxel gain (VI) and count them as views’ IG. Potthast and Sukhatme\textsuperscript{19} and Daudelin and Campbell\textsuperscript{20} used the relationship between various types of voxels to predict the occupancy probability of unobserved voxels, which makes the calculated IG of the view more reliable. The combined methods use both the surface information and the voxel information of the object, and the effect of these methods is better than the first two methods since more detection information of the object can be used. However, there are few studies with this method at present. Kriegel et al.\textsuperscript{21} proposed a combined algorithm that use mesh model to determine the candidate views and then calculate the IG of each candidate views. The view with the maximum IG is selected as the NBV.

Wei et al.\textsuperscript{22} represented a semi-model-based method which is suitable for 3-D building reconstruction with publicly available 2-D map data and estimated building height information. They estimate the rough shape of the object offline and plan the motion path of the sensor based on predicted model. Surface prediction based on prior information is more reliable than empirical surface prediction, so the view gain calculated by this method is more accurate. However, such methods require a higher amount of prior information, and the geometric characteristics of the target need to be fully known.

Method overview

This method is proposed for modeling geometry with relatively regular shapes, which have partial prior information before the active reconstruction. The priori information of the object includes the general features of the category to which the object belongs and the available features specific to the object. A multi-module cooperative system is designed which is suitable for the semi-model-based active object reconstruction mentioned above. It contains inter-process communication interfaces between modules. Modules of our system and their communication relationship are shown in Figure 1.

- **Data Receiver module** receives the sensor scheduling instruction, acquires the sensing data at a new view, and outputs it in the form of pointclouds or images.
- **Model Updater module** accepts perceptual data or manually tagged data to update the volumetric model, which is used for the IG calculation of views.
- **IG Calculator module** uses the information contained in the map to calculate views’ IG in a certain form.
- **NBV Planner module** infers with a priori information library in the rough shape estimation phase and evaluates each candidate view in the complete object reconstruction phase to determine the NBV.
- **Prior Information Library Creator module** determines a reasonable expression of the prior information, so that the knowledge library can be expanded freely.
- **Rough Shape Estimator module** guides the inference of the rough shape determination based on the prior information library and the perceptual data.
- **Object Surface Predictor module** generates point-clouds of the predicted surface and uses it to update the predicted occupied voxels in the model updater module.

The structure of this active object reconstruction method is summarized as follows:

1. Adjust the sensor to the specified pose and collect data.
2. Update the map with sensor data.
3. Determine the rough shape of the target based on the perceptual data and the prior information library.
4. If current data are not enough to determine the feature of object, call the view planning rule in the knowledge library and go to step 7. Otherwise, go to step 5.
5. Generate the predicted points that can be used to mark the predicted occupied voxels in the map.
6. Calculate the IG of candidate views.
7. Determine the NBV. End the algorithm when the termination condition is reached. Otherwise, go to step 1.

Library establishment and rough shape estimation

**Prior knowledge library establishment**

The prior information of the object includes feature information and relative position information. The feature
information describes the characteristics of a certain structure of the object, which can be used to determine the rough shape of the object. The relative position feature describes the positional relationship between different parts of the object and it can guide view planning during the stage of rough shape estimation. To express the priori information of the object in a unified structure to facilitate the call of it, we establish an extensible library, which can be initialized based on the known characteristics of the object before each new active vision task is performed.

The representation of rules in production system is drawn on and the form of its rule "premise → action" is improved by adding an additional function item. The expression of our rules is as follows

\[ I_k : \{ C_k \rightarrow R_k(\text{Proof}_k(), \text{Cal}_k(), \text{Move}_k())\} \]

Here the premise \( C_k \) and conclusion \( R_k \) refer to the geometric elements of the object, and \( R_k \) can be derived by \( C_k \). The rule is triggered when its premise is known and its conclusion is unknown. When its premise is an empty set, it indicates that this rule represents feature information of the object, and the latter element can be calculated by the perceptual data. Otherwise, it means that this rule represents the relative position between two areas. \( \text{Proof}_k() \) represents an activation function of a rule that judges whether it can reason about the relationship between the premise and the conclusion based on the observed data, and the rule can be activated if it returns true. \( \text{Cal}_k() \) is an attribute calculation function which is used to calculate the attributes of elements in \( R_k \) when the rule is activated. \( \text{Move}_k() \) is a movement function for NBV planning.

**Rough shape estimation algorithm and view planning at this stage**

An inference algorithm is proposed which cannot be constrained by specific perceptual tasks. It uses sensors’ detected information and the priori information library as a repository to determine the rough shape of the object. The outline of the inference algorithm is shown in Figure 2.

The **Assumption Set** stores elements that can fully express the rough shape of the object and each element corresponding to the premise or conclusion of the **Rules**. It means that the rough shape of the object is not determined when there is an unknown element in the **Assumption Set**. We only consider the unknown elements and compare them with the conclusion of each rule to determine whether the rule is triggered. If so, the activation function of the rule is execute. We run the attribute calculation function if the rule is activated and store its characters into the **Feature Set**. When all rules cannot be triggered or activated, it indicates that we cannot infer the rough shape of the object at the **Current View** and we need to plan the NBV to obtain more features. If all elements in the **Assumption Set** are calculated, then we generate points of the object’s predicted surface based on these features and update the volumetric map which is used in the second stage of active reconstruction.

The view planning algorithm at this stage aims to find more features of the object. The movement of the sensor is determined by relative positions between two geometry elements of the object, which is included in rules of the priori information library. The rules are called for NBV planning, and the algorithm is concluded as Algorithm 1. IG at this stage is calculated using off-the-shelf methods.

**Algorithm 1.** The view planning algorithm at first stage.

```plaintext
for each each rule in the priori information library do
  if it is triggered then
    Calculate the next view position and put it into the candidate view set
  end if
end for
if candidate view set is not empty then
  Calculate the distances between current view and candidate views, choose the one who gets the minimum distance as the NBV.
else
  Calculate the IG of the initial candidate views, choose the one who get the maximum as the NBV
end if
```

**IG calculation**

IG is employed for view planning in the complete reconstruction stage. Since our goal is to get more surface of the object, IG here represents the gain of surface coverage. When the pointclouds of the object’s predicted surface is generated, cells corresponding to each point...
are marked as predicted occupied voxels in the volumetric model space, which is expected to be closest to the actual target surface.

The set of candidate views is denoted as $V$ in which view $v \in V$ is generated by a certain method. A set of voxels $C$ are traversed by the ray $r$ that belongs to the set of rays $R_v$ which cast through the map on view $v$. We calculate view’s IG $G$ by collecting the gain of each voxel $c$ passed by rays in $R_v$ (see equation (1))

$$G(v) = \sum_{r \in R_v} \sum_{c \in C} I(c)$$

Voxel’s IG $I$ is defined like equation (2). We consider that the first cell passed through by a ray has the IG of 1 and other cells have the value of 0.

$$I(c) = \begin{cases} 1 & \text{if } c \text{ is the first voxel traversed by the ray} \\ 0 & \text{otherwise} \end{cases}$$

View’s IG calculated in this way indicates the area of predicted surfaces without observation that the view can acquire, and view $v_n$ with the largest IG is selected as the NBV, as shown in equation (3)

$$v_n = \arg\max_v G(v)$$

As shown in Figure 3, our volumetric map is composed of voxels of the same size. Green cells are free, black cells are occupied, gray cells are unknown, pink cells are predicted occupied, and dark gray cells are predicted interior. Yellow rectangles mean a sensor, and red lines indicate rays of the sensor. Since occupied voxels represent the surface of object, voxels that are within sensor’s measurement range and in front of the occupied voxels in the direction of a ray are taken into consideration.

**Experiment results**

To demonstrate the performance of our proposed algorithm, this active reconstruction algorithm is utilized on multiple objects in a simulated scene. Our simulation environment is based on robot operating system\textsuperscript{23} and Gazebo, and we use Octomap\textsuperscript{24} that has a resolution of 0.01 m as the volumetric map. The reconstructed objects are 3-D polygonal prism or pyramid models generated by 3DsMax, including a 10 prism, a 20 prism, a 50 prism, and a hexagonal pyramid. The initial models are shown in Figure 4(a) to (d). A simulated Kinect with a measurement range of 1.5 m is employed as the sensing tool, which can return image and pointclouds data. The received data are used to update the map and calculate the rough shape of object. Candidate views are obtained by uniformly spreading points on a spherical surface before reconstruction, and they point to the center of the sphere. The termination condition of the algorithm is to move the sensor 20 times. The reconstruction results of the objects are shown in Figure 4(e) to (h).

To measure the speed of algorithms in obtaining object information, we calculate the object’s surface coverage after each view moving and observe its growth rate. We discretize a series of points of the initial model and compare them with the pointclouds obtained during the reconstruction process. If there is a point whose distance from the initial point is less than the threshold, then the initial point is considered to have been detected. We define the threshold as 0.005 m. The surface coverage is expressed as a percentage of the detected points to all initial points. The number of views required for the complete reconstruction is one of the important metrics of the active reconstruction algorithm, which embodies the consumption of sensors and has a clear physical meaning. The fewer the number of views used for reconstruction, the better the performance of the algorithm. Due to factors such as sensor error, the surface coverage may not always reach 1. We assume that the modeling is completely reconstructed when the object’s surface coverage reaches 99%.

We compare our algorithm with the other seven algorithms. They all belong to the search-based methods, and their difference lies in the way in which the view’s IG is calculated. Voxel’s IGs brought forward by Delmerico et al.\textsuperscript{25} include Occlusion Aware VI, Unobserved Voxel VI, Rear Side Voxel VI, Rear Side Entropy VI, and Proximity Count VI. Vasquez-Gomez et al.\textsuperscript{26} proposed an IG calculation method which can be called Area Factor VI. Another method proposed by Kriegel et al.\textsuperscript{21} is Average Entropy VI.

Different algorithms are used on various models in the same environment, and the corresponding surface coverage is shown in Figure 5. It can be seen from the figure that on some objects, the speed of coverage improvement at the first few views by our method is defeated by some other...
methods, and this might be caused by the goal of estimating the rough shape of the object at the first stage that conflicts with the purpose of obtaining more surface areas. But the rough shape estimation at this stage will provide more help for view planning at the next stage. Table 1 shows the number of views required for the reconstruction of each method. According to the table, it can be seen that our algorithm basically needs the minimum views, which...
proves that the loss of surface coverage growth rate at the first few views is positive for overall reconstruction.

Conclusion

In this work, a two-step active object reconstruction method based on the partial prior information of the object is proposed, where the object has relatively regular shape but its specific characteristics are unknown. A fixed form of prior knowledge library was created and it is filled before reconstruction according to the specific tasks. An inference engine based on the prior knowledge library was proposed, which will not be affected by the specific content of the prior knowledge library. It is used to estimate the object’s rough shape and determine the NBV at the first stage of active reconstruction. In addition, a new IG calculation method based on the rough shape of object is proposed to plan the NBV at the second stage.

Our algorithm is evaluated and compared with seven other methods with simulated experiments. Our experiments isolated the performance of each method, without effects from environmental factors. The results proved that our algorithm can complete the reconstruction in a few views. By comparing the results of each algorithm in the speed of surface coverage improvement and the number of views required for reconstruction, it can be found that our algorithm performs generally the same or better than other algorithms.

However, it performs mediocre when faced with the simple geometric objects, and this defect needs to be improved. Combining the optimization goal of estimating the rough shape with obtaining more surfaces at the first stage might be an effective solution, which will be further researched in future work. In addition, a more generalized prior information library to make full use of the various kinds of prior information of objects is worth to be established as a next step. Besides, this algorithm is applied in the simulation environment at present, and in the next step, it will be utilized to the modeling of buildings in the real environment.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China under grant no. U1713216.

ORCID iD

Yanzi Kong https://orcid.org/0000-0003-0780-6966

References

1. Koniar D, Hargáš L, Loncová Z, et al. Machine vision application in animal trajectory tracking. Comput Methods Programs Biomed 2016; 127: 258–272.
2. Koniar D, Hargáš L, Loncová Z, et al. Visual system-based object tracking using image segmentation for biomedical applications. Electr Eng 2017; 99(4): 1349–1366.
3. Haffner O and Duchiň F. Making a map for mobile robot using laser rangefinder. In: 23rd international conference on robotics in Alpe-Adria-Danube Region (RAAD), Smolenice, Slovakia, 3–4 September 2014, pp. 1–7.
4. Tőgyessy M, Dekan M, Duchiň F, et al. Foundations of visual linear human–robot interaction via pointing gesture navigation. Int J Soc Robot 2017; 9(4): 509–523.
5. Turygin Y, Bozek P, Nikitin Y, et al. Enhancing the reliability of mobile robots control process via reverse validation. Int J Adv Robot Syst 2016; 13(6): 1–8.
6. Lozhkin A, Abramov I, Bozek P, et al. The issue of calculating elliptic trajectories. Manuf Technol 2014; 14(4): 561–566.
7. Grau V, Mewes AUJ, Alcaniz M, et al. Improved watershed transform for medical image segmentation using prior information. IEEE Trans Med Imaging 2004; 23(4): 447–458.
8. Connelly C. The determination of next best views. In: IEEE International conference on robotics and automation, St. Louis, MO, USA, 25–28 March 1985, pp. 432–435. USA: IEEE.
9. Aloimonos J, Weiss I, and Bandyopadhyay A. Active vision. Int J Comput Vis 1988; 14(4): 333–356.
10. Scott W, Roth G, and Rivest JF. View planning for automated three-dimensional object reconstruction and inspection. ACM Comput Surv 2003; 35(1): 64–96.
11. Chen SY, Li YF, and Kwok NM. Active vision in robotic systems: a survey of recent developments. Int J Robot Res 2011; 30(11): 1343–1377.
12. Mavrinac A, Chen X, and Alarcon-Herrera JL. Semi-automatic model-based view planning for active triangulation

Table 1. Number of views required for complete reconstruction.

|          | Occlusion Aware VI | Unobserved Voxel VI | Rear Side Voxel VI | Rear Side Entropy VI | Proximity Count VI | Area Factor VI | Average Entropy VI | Ours |
|----------|-------------------|---------------------|--------------------|----------------------|--------------------|----------------|--------------------|------|
| Ten prism| 12                | 8                   | 8                  | 9                    | 6                  | 20+           | 5                  | 6    |
| Twenty prism| 18               | 20                  | 18                 | 17                   | 20                 | 20+           | 13                 | 11   |
| Fifty prism| 20+               | 20+                 | 14                 | 19                   | 20+                | 20+           | 10                 | 8    |
| Hexagonal pyramid | 5            | 4                   | 5                  | 6                    | 4                  | 4             | 6                  | 4    |

The bold values in table means the minimum views needed for each object.
3-D inspection systems. *IEEE/ASME Trans Mechtron* 2015; 20(2): 799–811.

13. Munkelt C, Breitbarth A, and Notni G. Multi-view planning for simultaneous coverage and accuracy optimization. In: BMVA. Aberystwyth, UK, September 2010, pp. 1–11.

14. Wakisak E, Kanai S, and Date H. Model-based next-best-view planning of terrestrial laser scanner for HVAC facility renovation. *Comput Aided Des Appl* 2018; 15(3): 353–366.

15. Chen SY and Li YF. Vision sensor planning for 3-D model acquisition. *IEEE Trans Syst Man Cybern B* 2005; 35(5): 894–904.

16. Li YF and Liu ZG. Information entropy-based viewpoint planning for 3-D object reconstruction. *IEEE Trans Robot* 2005; 21(3): 324–337.

17. Isler S, Sabzevari R, Delmerico J, et al. An IG formulation for active volumetric 3D reconstruction. In: *IEEE international conference on robotics and automation*. Stockholm, Sweden, May 2016, pp. 3477–3484. Sweden: IEEE.

18. Vasquez-Gomez JI,Sucar LE, Murrieta-Cid R, et al. Volumetric next-best-view planning for 3D object reconstruction with positioning error. *Int J Adv Robot Syst* 2014; 11(10): 159–171.

19. Potthast C and Sukhatme GS. A probabilistic framework for next best view estimation in a cluttered environment. *J Vis Commun Image R* 2014; 25(1): 148–164.

20. Daudelin J and Campbell M. An adaptable, probabilistic, next-best view algorithm for reconstruction of unknown 3-D objects. *IEEE Robot Autom Lett* 2017; 2(3): 1540–1547.

21. Kriegel S, Rink C, Bodenmüller T, et al. Efficient next-best-scan planning for autonomous 3D surface reconstruction of unknown objects. *J Real-Time Image Process* 2015; 10(4): 611–631.

22. Wei J, Polden J, Tao P, et al. View planning for 3D shape reconstruction of buildings with unmanned aerial vehicles. In: *International conference on control, automation, robotics and vision (ICARCV)*. Phuket, Thailand, 13–15 November 2016, pp. 1–6. Thailand: IEEE.

23. Quigley M, Conley K, Gerkey B, et al. ROS: an open-source robot operating system. In: *ICRA workshop on open source software*. Kobe, Japan, 17 May 2009, pp. 1–6. Japan: IEEE.

24. Hornung A, Wurm KM, Bennewitz M, et al. OctoMap: an efficient probabilistic 3D mapping framework based on octrees. *Auton Robot* 2013; 34(3): 189–206.

25. Delmerico J, Isler S, Sabzevari R, et al. A comparison of volumetric IG metrics for active 3D object reconstruction. *Auton Robot* 2018; 42(2): 197–208.

26. Vasquez-Gomez JI, Sucar LE, and Murrieta-Cid R. View planning for 3D object reconstruction with a mobile manipulator robot. In: *IEEE/RSJ international conference on intelligent robots and systems*. Chicago, IL, USA, 14–18 September 2014, pp. 4227–4233. USA: IEEE.