



# Object pose estimation for robot loading in accommodation space using alpha-shape algorithm

Qingda Guo<sup>1</sup> · Lulu Tang<sup>2</sup> · Jianchi Zhang<sup>1</sup>

Accepted: 3 September 2021 / Published online: 21 September 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

## Abstract

Robots with visual sensors have been used in various goods logistics, such as bin picking or uploading. However, there are more and more demands for the automatic blanking and loading, and it is necessary to solve the problem of object pose estimation in changing accommodation space. This paper proposes a method for pose estimation in the accommodation space using alpha-shape algorithm and improved fruit fly optimization algorithm (FOA). The alpha-shape volume variety of object and measured space is set to the objective function, and the pose variety of object is set to six variables of improved FOA. The experiments were performed by setting parameters of improved FOA and considering the four space types represented the common accommodation shapes. Compared with previous work using convex hull, the new study using alpha-shape algorithm not only keeps the object in the accommodation space, but also maintains the object pose which is at the bottom of the space and can meet the practical requirement of object placement by robot arms.

**Keywords** Robot · Pose estimation · Alpha-shape algorithm · Fruit fly optimization · Accommodation space

## 1 Introduction

With the increase in labor cost and the automation demand of production line equipment, industrial robots with stereo vision are applied to product processing, especially for bin picking and workpiece loading, such as the picking competition held by Amazon from the box or on the shelf. However, there are more and more demands for the automatic blanking and placement of objects in the “reversed picking process.” Limited by the visual perception technology of industrial robots, the packing process is usually set by manual procedure or based on artificial prior knowledge. For example, objects are packaged into

simplified geometric shapes (cubes, cylinders, etc.) and then loaded into the boxes with known shapes and fixed positions. For some industrial robots with complex shape and elastic objects, robot engineers often set up the procedure of goods placement. Wynright company of the USA (Criswell 2014) built a mobile robot system for automatic loading of tires in containers, but there will still be unexpected space changes in the container body and thus remaining space (accommodation space). The key problem of automatic loading in an irregularly changing space is to obtain the pose of a target object in the remaining accommodation space, which eventually contains at least one object.

The effective pose matrix of the object in the accommodation space is the last crucial step in the process of intelligent robot’s picking, running (obstacle avoidance) and placing. Automatic loading (Guo et al. 2019) using industrial robot needs the hand–eye matrix (1), pose matrix (2) between the camera coordinate system and workpiece coordinate system of reference model, pose matrix (3) between the reference model and the target model and pose matrix (4) between the target object (or the reference model (5)) and the accommodation space, as shown in Fig. 1a. In the past few years, object recognition of robot in

---

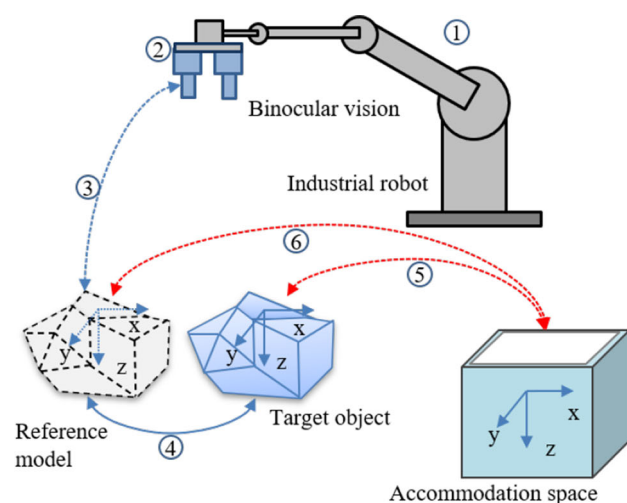
✉ Qingda Guo  
guoqd@scut.edu.cn

Lulu Tang  
lulu.tang@connect.um.edu.mo

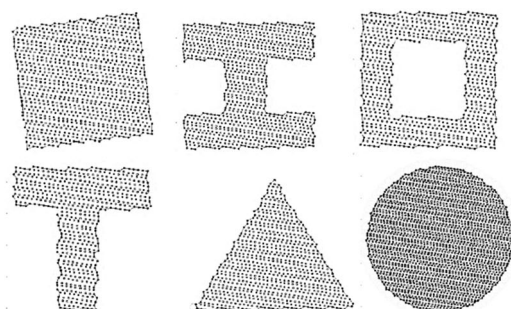
Jianchi Zhang  
msj.c.zhang@scut.edu.cn

<sup>1</sup> School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China

<sup>2</sup> State Key Laboratory of Internet of Things for Smart City, University of Macau, San Francisco, CA, USA



(a) Structure diagram of automatic loading using industrial robots in the remaining accommodation space [2]



(b) The applications of alpha-shape algorithm [13]

**Fig. 1** Structure diagram of the system and the applications. **a** Structure diagram of automatic loading using industrial robots in the remaining accommodation space (Guo et al. 2019). **b** The applications of alpha-shape algorithm (Wang and Chen 2019)

bin picking has been a hot topic in the scientific research and robot companies. Many algorithms have been proposed in the aspect of 6-degree of freedom (DOF) pose estimation, such as the application of Fast Point Feature Histograms (FPFH), Signature of Histograms of Orientations (SHOT), LineMod algorithm with template (Hinterstoisser et al. 2012), Point Pair Feature (PPF) with voting mechanism (Drost et al. 2010), DenseFusion (Wang et al. 2019), Pose-CNN based on regression (Xiang et al. 2017) and ssd-6d (Kehl et al. 2017).

The end-to-end trajectory of robot is also a hotspot research, which focuses on the trajectory planning based on optimal time or minimum energy. These planning algorithms mostly establish the dynamic and kinematic constraints of robot, using intelligent optimization algorithm to achieve trajectory acquisition or apply sensors to generate Octomap or three-dimensional (3D) data environment to

establish space obstacle avoidance trajectory. Robot is usually based on the pre-setting pose matrix when it moves to the place where objects can be placed. Human beings have great flexibility in the operation of object placement, such as choosing a reasonable place for the object according to the size of the object or the environment on the plane, and packing in a limited space, which can be best placed according to different objects.

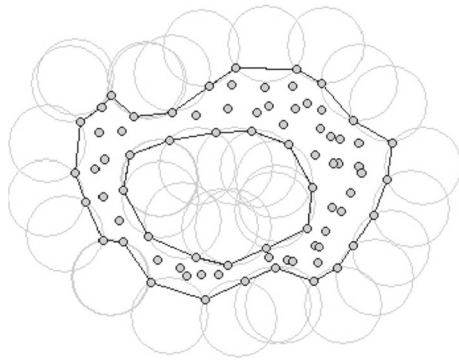
Alpha-shape algorithm is used to reconstruct object surface from an unorganized point cloud. The method was proposed in 2D points by Edelsbrunner and was then extended to 3D points (Edelsbrunner et al. 2003; Edelsbrunner and Mücke 1994). Compared with convex hull, alpha shape can be used to reconstruct object shape of nonconvex body, as shown in Fig. 1b. Alpha shape was widely used in 3D object shape. Zhu et al. (2008) proposed a novel approach for tree crown reconstruction based on an improvement of alpha-shape modeling, where the data are points unevenly distributed in a volume rather than on a surface only. Lou et al. (2013) used alpha shape to extract topographical features from engineering surfaces and found that the alpha-shape method was more efficient in performance for large structuring elements. Santos et al. (2019) proposed an adaptive method which estimates a local parameter for each edge based on local point spacing and used the method to extract the building roof boundaries from LiDAR Data.

Based on previous study (Guo et al. 2019), this paper still aims to solve the problem of object pose estimation in the accommodation space assuming that accommodation space can hold one target object. The alpha-shape algorithm and improved FOA (Pan 2012) are used to determine the object's final state given a pose matrix. Firstly, the alpha-shape algorithms are introduced, and then, the alpha-shape volume variety of object and measured space is set to the objective function and the pose variety of object is set to six variables of improved FOA. Next, in the simulation experiments we at first present pose estimation process in cube space and obtain the convergence curves of four different spaces (namely the cube, hemisphere, cylindrical and triangular prism) and the object's final pose in the four spaces. Finally, the reliability of the proposed method is verified experimentally, comparisons with previous work are given, and conclusions are drawn.

## 2 Proposed method

### 2.1 Alpha-shape algorithm

Alpha-shape algorithm can reconstruct geometry from a set of discrete points, as shown in Fig. 2. A circle with radius value  $\alpha$  is used to scroll around the point set  $S$ . After



**Fig. 2** Alpha shape of the planar point set

traversing all the points, the inner and outer contours of  $S$  can be obtained. When the radius  $\alpha$  is large, the circle rolls outside  $S$ , and the trace of the external roll is the boundary contour of the point set. When  $\alpha$  is small, the circle will roll to the interior of  $S$ . When  $\alpha$  is small enough, every point in  $S$  is a boundary point. The value of circle radius  $\alpha$  is closely related to the fineness of the detected points contour. When the radius is relatively small, the detected contour is relatively fine, and when the radius is relatively large, the detected contour is slightly rough.

## 2.2 Pose estimation algorithm

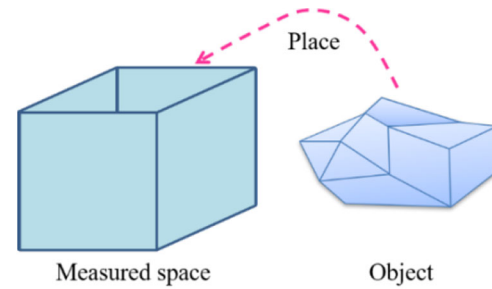
When an industrial robot picks, runs, and places the target object in the accommodation space, the shape of accommodation space generally can be generally convex. The swarm intelligence optimization algorithm usually requires an objective function to determine the direction of iterative optimization. Compared with previous work (Guo et al. 2019), we just use the volume of alpha shape to establish an objective function of FOA.

In the process of placing the object in the accommodation space, there are two groups of point clouds (the object and the accommodation space), in Fig. 3a–c. The two groups of points are combined into one group, and the change of relative position can be expressed by the whole volume value  $V_t$ . The whole volume value  $V_t$  of the optimal position is minimal, as the object is placed at the bottom of the accommodation space and there is no space between the object and the bottom of the accommodation space.

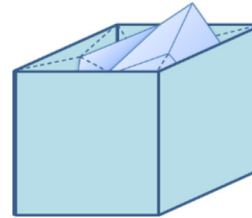
As the object is placed at the bottom of the accommodation space, the combined point cloud is concave. The whole volume value  $V_t$  can be obtained by tuning  $\alpha$  value of alpha-shape algorithm. The fitness function of FOA is expressed by:

$$S_{\min} = V_t. \quad (1)$$

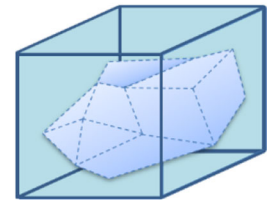
Based on previous work (Guo et al. 2019), the target rotation is regarded as a separate rotation along each axis,



**(a)** Measured accommodation space and a target object



**(b)** Initial placement



**(c)** Final placing result

**Fig. 3** The process of object placing into the measured space

and the posture adjustment of the object is simplified into three parameter variables, whose rotation matrixes are expressed as:

$$\mathbf{R}_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix}, \quad (2)$$

$$\mathbf{R}_y(\beta) = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix}, \quad (3)$$

$$\mathbf{R}_z(\gamma) = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4)$$

where  $\theta$ ,  $\beta$  and  $\gamma$  are the rotation parameters of  $x$ -,  $y$ - and  $z$ -axes, respectively, and  $\mathbf{R}_x(\theta)$ ,  $\mathbf{R}_y(\beta)$  and  $\mathbf{R}_z(\gamma)$  are the corresponding rotation matrices, respectively.

To overcome local optimum, we still use multiple individuals to search for a global optimum, as shown in Fig. 4. We start by defining the point cloud of the target object as  $O(x, y, z)$  and the point cloud of the measured accommodation space as  $P(x, y, z)$ .

(1) Initialize the FOA parameters:

$$O_g(\bar{x}_m, \bar{y}_m, \bar{z}_m) = \sum_{i=1}^m O_i(x_i, y_i, z_i)/m, \quad (5)$$

$$P_g(\bar{x}_n, \bar{y}_n, \bar{z}_n) = \sum_{i=1}^n P_i(x_i, y_i, z_i)/n, \quad (6)$$

$$[O^1, 1]^T = \begin{bmatrix} \mathbf{E} & \mathbf{O}_g - P_g \\ 0 & 1 \end{bmatrix} \cdot [O, 1]^T, \quad (7)$$

$$V^t = \text{alphaShape}([O^1; P_g]), \quad (8)$$





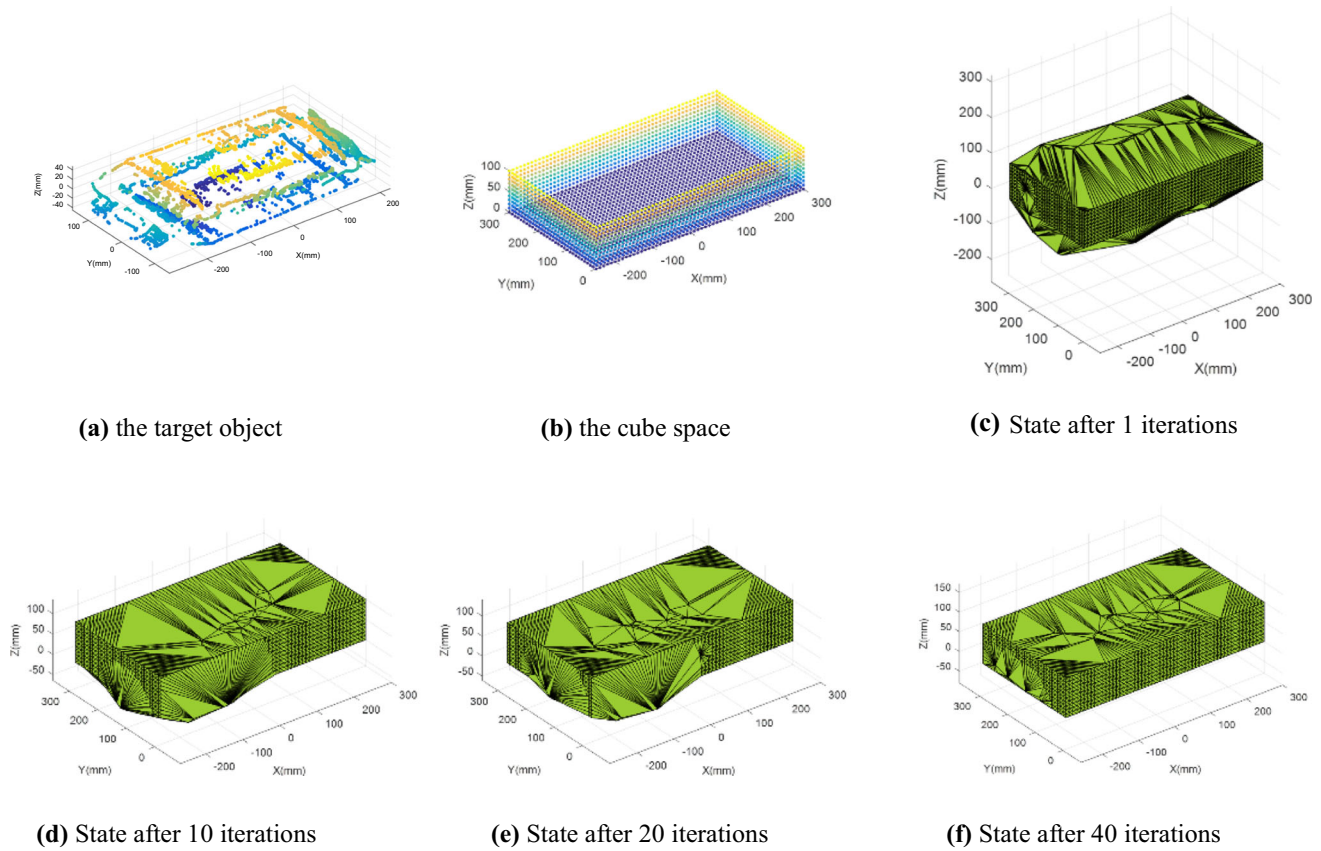


Fig. 6 Iteration process and its final results

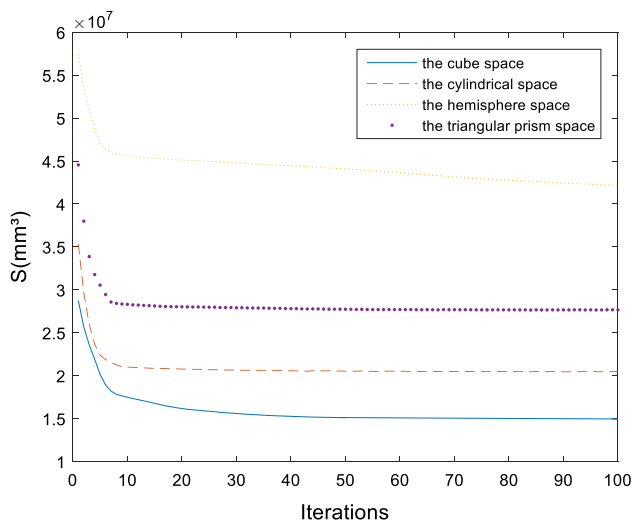


Fig. 7 The iterative convergence process

$$[S\_s S\_index] = \text{sort}(S). \quad (19)$$

where  $S\_s$  is the volume value after sorting the groups and  $S\_index$  is index of the corresponding individual.

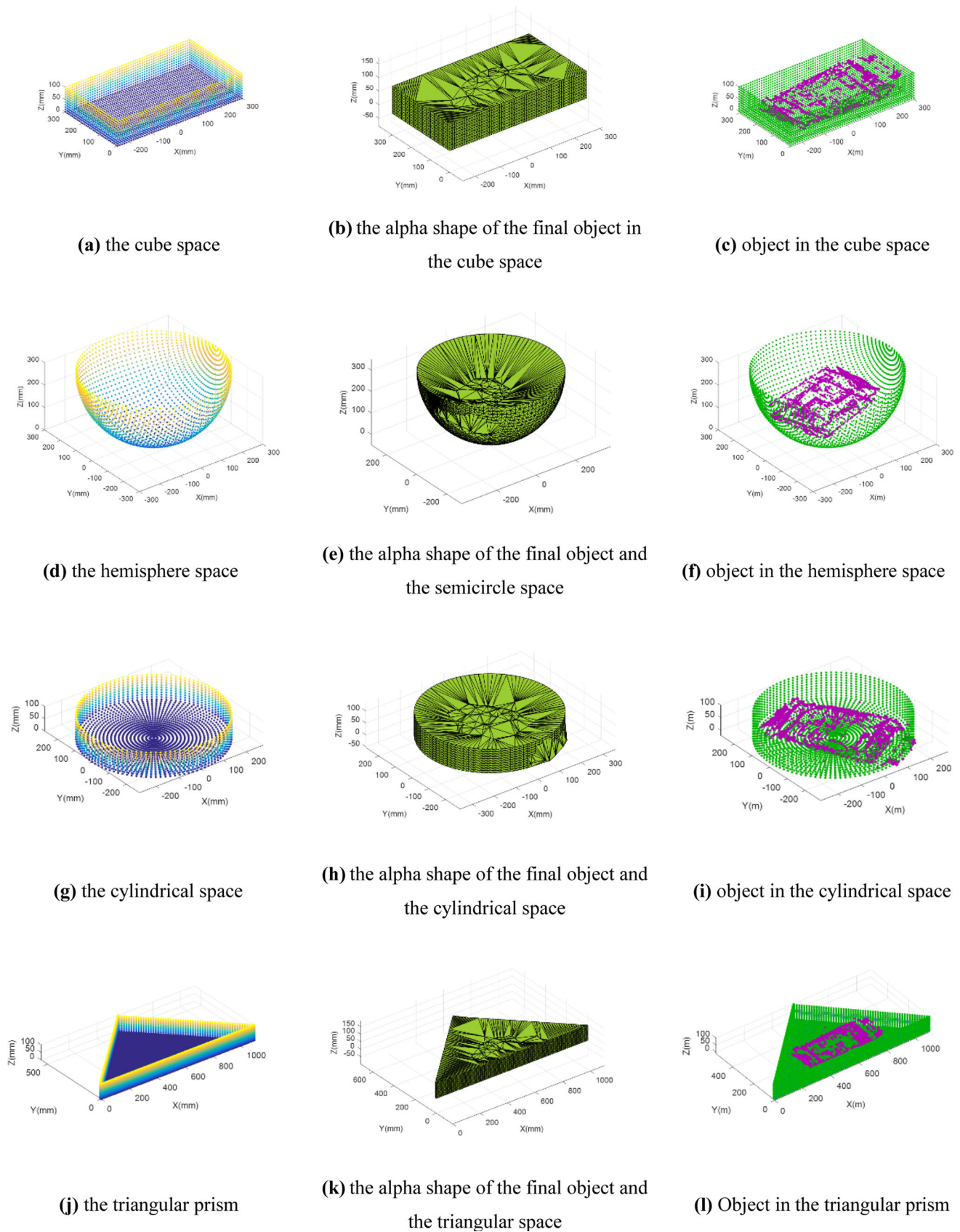
(6) Half individuals with smaller concentration values are selected and marked again according to the concentration values:

$$\begin{cases} X_j = X(S\_index) \\ Y_j = Y(S\_index) \\ Z_j = Z(S\_index) \end{cases}, \quad (20)$$

$$\begin{cases} \alpha_j = \alpha(S\_index) \\ \beta_j = \beta(S\_index) \\ \gamma_j = \gamma(S\_index) \end{cases}, \quad (21)$$

where the range of  $S\_index$  is determined by the number of individuals in half population,  $(X_j, Y_j, Z_j)$  denote new position,  $(\alpha_j, \beta_j, \gamma_j)$  denote new posture, and subscript  $j$  is index of individual.

(7) After step (6), return to step (3), and then to step (4); random direction and location are updated to obtain the taste concentration of a new half individuals. The pre-selected half individuals and the new half individuals are combined, and then, all individuals are sorted by the taste concentration according to step (5). The algorithm will stop the iterative optimization loop until the end condition is satisfied.



**Fig. 8** The designed types of accommodation space

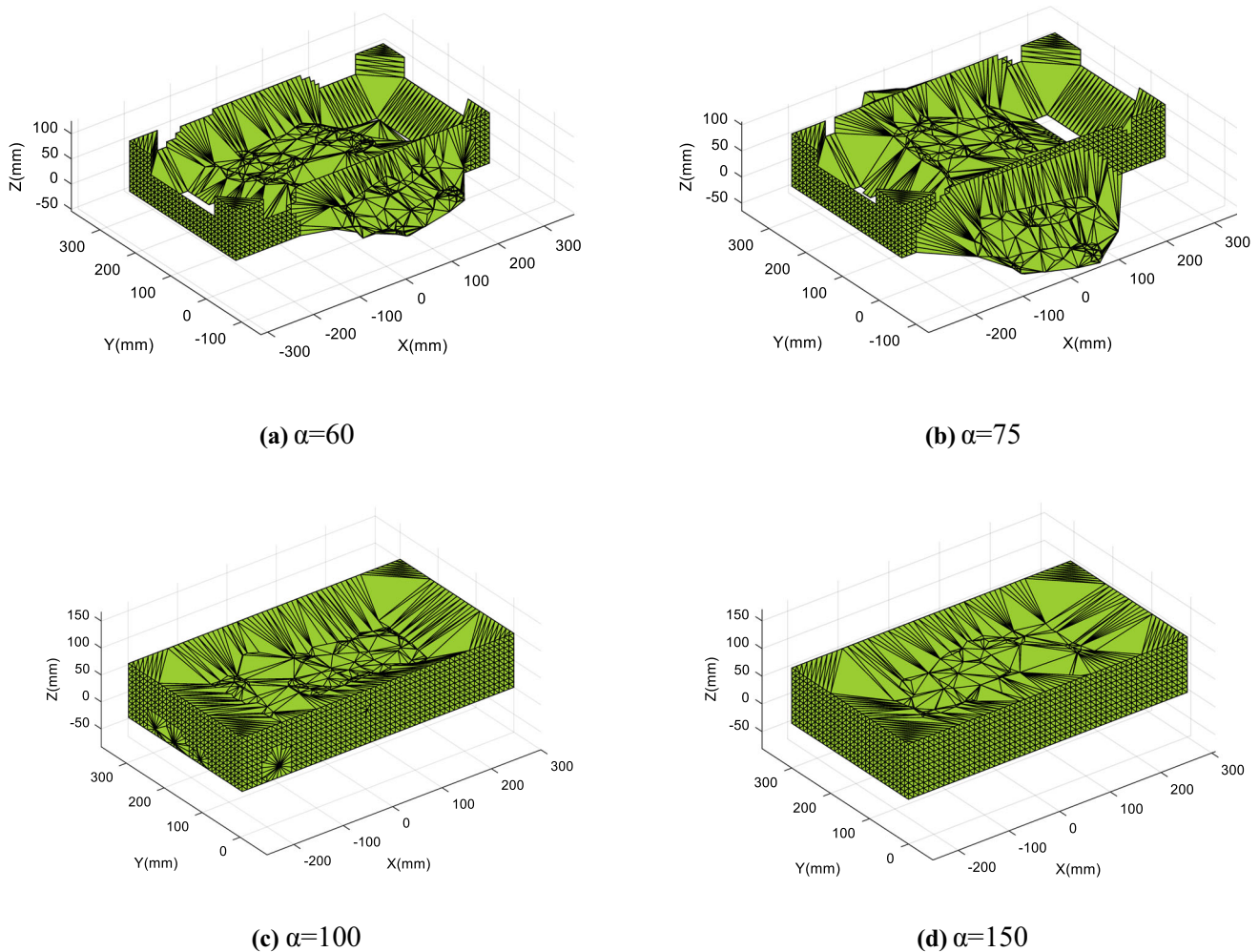


Fig. 9 The alpha shape with different parameter  $\alpha$

### 3 Experimental results

#### 3.1 Obtaining the target object

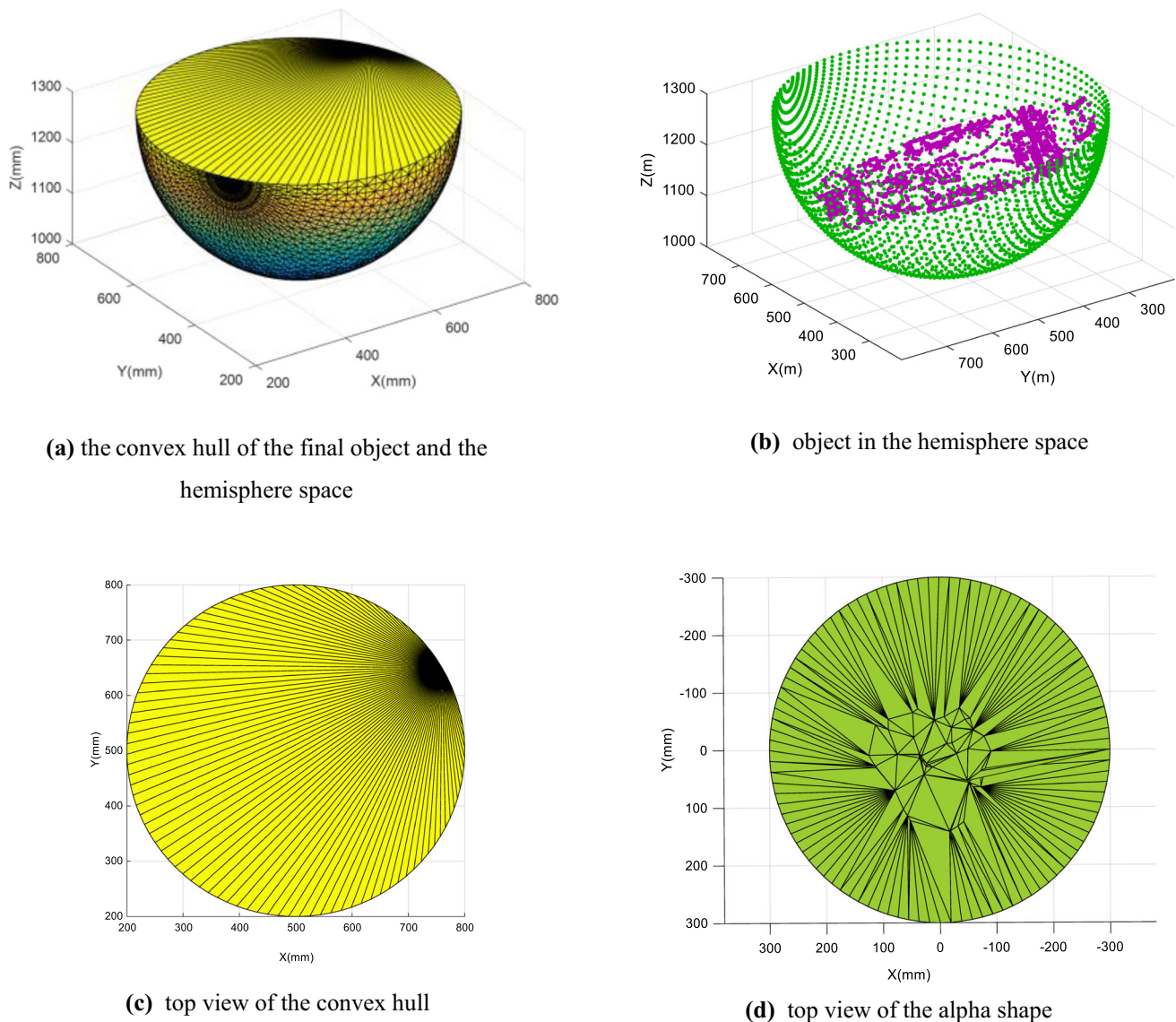
In order to obtain the sparse point cloud of the target, the object was placed in the common field of view of the left and right cameras, and the object images of the left and right cameras were obtained, as shown in Fig. 5a. According to the point cloud acquisition process, Otsu threshold segmentation was performed on the object image of the left camera; subsequently, the background was removed, and region of the object was retained, as shown in Fig. 5b. Harris feature points were detected on the left image after threshold segmentation, as shown in Fig. 5c. The feature points of the right image were tracked by the obtained left image feature points, and the positions of features point pairs were obtained using KLT optical flow (Shi and Tomasi 1994), as shown in Fig. 5d. According to the corresponding feature point positions of left and right images and camera calibrated parameters, the top point

clouds in the left camera coordinate system were calculated and finally rotated and translated into the world coordinate system, as shown in Fig. 5e. The bottom point cloud was obtained using the same method, and the top point cloud and the bottom point cloud were combined to a complete point cloud of the target object based on the world coordinate system.

#### 3.2 Parameters setting result analysis

The initial parameters of the improved FOA are population size, number of iterations and random direction and position parameters ( $a_1 \sim a_6$  and  $b_1 \sim b_6$ ). In the experiments, the number of iterations is set to be 100 and the radius  $\alpha$  is set to be 250. The range of random position variation was limited to  $(-1, 1)$ , and the range of random posture variation was limited to  $(-0.05\pi, 0.05\pi)$ ; thus,  $a_1 \sim a_3$  and  $b_1 \sim b_3$  are set to be 2 and 1, respectively, and  $a_4 \sim a_6$  and  $b_4 \sim b_6$  are set to be  $0.1\pi$  and  $0.05\pi$ , respectively. Software environment is Win10 64bit and MATLAB





**Fig. 10** The object pose estimation using convex hull

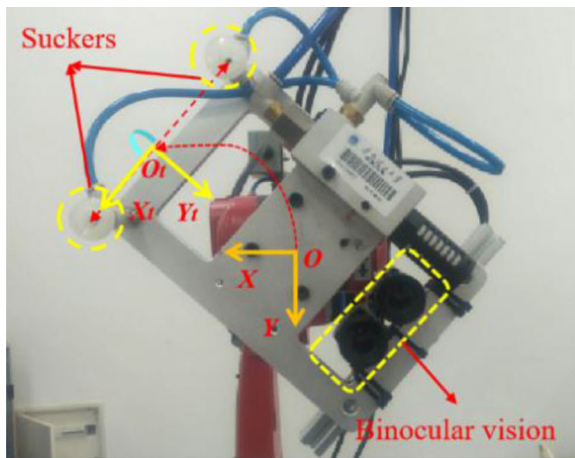
2015b, and the corresponding hardware configuration parameters were CPU i7-7700 and RAM 8 GB.

Based on previous study, the point cloud of the target object is obtained by the Harris feature point detection and KLT optical flow, in Fig. 6a. 3D accommodation space is set as a cuboid with the size of  $550 \times 100 \times 300 \text{ mm}^3$ , in Fig. 6b. The population number of fruit flies is 100. Figure 6c shows the initial state; Fig. 6d shows the 10th iteration; Fig. 6e shows the 20th iteration; and Fig. 6f shows that pose adjustment in the accommodation space is completed at the 40th iteration.

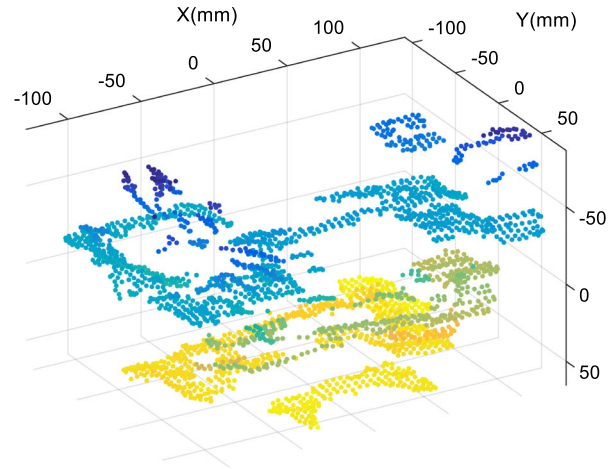
The space types are designed only to accommodate the object and the target object's final posture in those three kinds of space as shown in Fig. 8. The four space types,

namely the cube, hemisphere, cylindrical and triangular prism, represent different shapes of the common accommodation space. The proposed method can be applied to different convex shaped spaces. The iterative convergence curves of the four space types are plotted in Fig. 7. The four curves show that the object is basically placed in four spaces after iterating the 10th iteration. To test the effect of different value of parameter  $\alpha$  on the estimation pose, parameter  $\alpha$  values are set to be 60, 80, 100, 150, respectively. As shown in Fig. 9, there will be different results for setting different values of parameter  $\alpha$  and it is important to choose proper  $\alpha$  value.

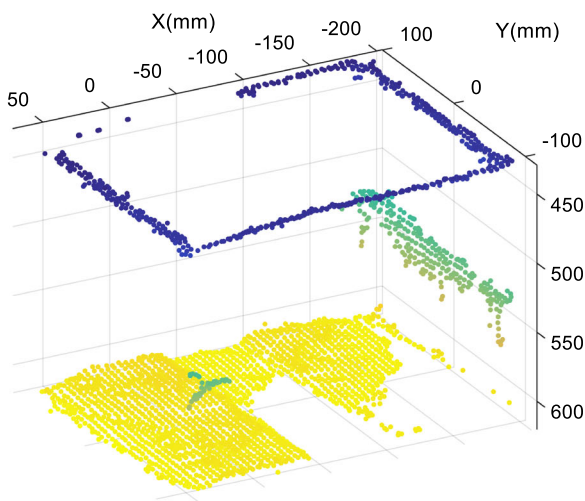




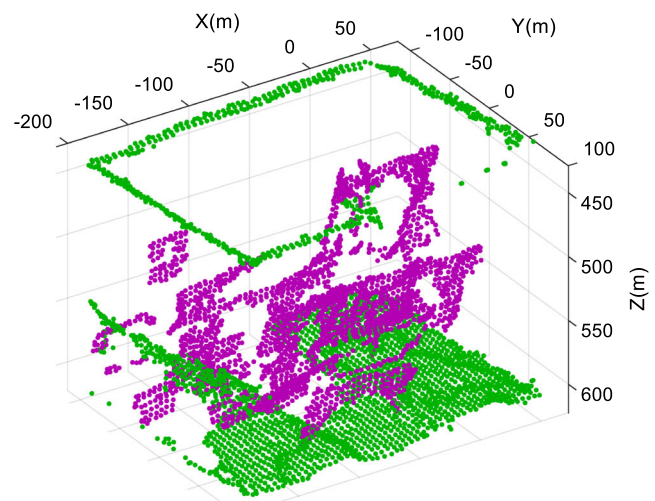
(a) The end-effector of the industrial robot



(b) Point cloud of the complex-shaped object



(c) Point cloud of the cube box



(d) The complex object pose in the cube box space

Fig. 11 Hardware platform test of the complex object

### 3.3 Method comparison

Our previous work uses convex hull to construct the point clouds of the object and the accommodation space. Object final pose in the hemisphere space is computed using previous method, as shown in Fig. 10. Comparing Figs. 8e–f, 10d with Fig. 10a–c, the result shows that alpha-shape algorithm is more suitable for the placement requirements of robot real scenes than the convex hull algorithm. Previous work only considers that the accommodation space meets the placement of the target object, but not gives the best final pose in practical application; our new study using alpha-shape algorithm not only keeps the object in the accommodation space, but also maintains the

object pose which is at the bottom of the space and can meet the actual robot placements requirements.

### 3.4 Hardware test

We used the SIASUN industrial robot (model: SR6C), binocular vision (or 3D vision sensor, model: Astra Mini) and vacuum sucker to establish the hardware test platform shown in Fig. 11a. A more complex-shaped object and a randomly cube box were considered, and their corresponding point clouds were derived using a 3D sensor, as shown in Fig. 11b, c. The final pose of the complex object in the box space was estimated using the proposed method, as shown in Fig. 11d.

## 4 Conclusion

In this paper, the method of combining alpha-shape algorithm and improved fruit fly optimization is proposed to estimate the object posing in a 3D accommodation space. The proposed method uses the alpha-shape algorithm to establish the objective function, which uses the whole volume change to adjust the object pose. To obtain the best individual with 6 degrees of freedom, the iteration strategy of improved FOA chooses better half individuals to produce next half individuals. The experiments were performed setting parameters of improved FOA and considering the four space types, and the obtained results show that proposed method can obtain the object pose in the common accommodation space. Compared with previous work using convex hull, the new study using alpha shape could keep the object pose at the bottom of the accommodation space without limitation of gravity, which could meet the actual robot stacking. For an actual object placed in the enclosed space, the future study would increase the constraints (such as elasticity and collision) and adjust the parameters (or adds other algorithms) to shorten the estimation time. In addition, the objective function depends on the whole volume change of alpha shape and parameter  $\alpha$  directly related to point density and the level of detail of the boundary. As the proposed method needs choosing an appropriate parameter  $\alpha$ , segmenting point clouds of the accommodation space and building 3D object, our future work will focus on using adaptive alpha-shape algorithm and obtaining the whole scene point clouds and the space point clouds segmentation.

**Acknowledgements** This work was supported by Key-Area Research and Development Program of Guangdong Province, China (2019B010155001). In addition, thanks to all the people who are fighting the COVID-19.

**Authors' contributions** Guo Qingda contributed to the conception of the study and wrote the manuscript; Tang Lulu contributed significantly to analysis and manuscript preparation; and Zhang Jianchi performed the experiment.

## Declarations

**Conflict of interest** All authors declare that they have no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

**Data Availability** All data generated or analyzed during this study are included in this article.

## References

- Criswell T (2014) Arlington, TX (US). Automated truck unloader for unloading unpacking product from trailers and containers. US, 2014/0205403 A1. 24 July 2014
- Drost B, Ulrich M, Navab N et al (2010) Model globally, match locally: efficient and robust 3D object recognition. In: IEEE conference on computer vision and pattern recognition
- Edelsbrunner H, Mücke EP (1994) Three dimensional alpha shapes. *ACM Trans Graph* 13(1):43–72
- Edelsbrunner H, Kirkpatrick D, Seidel R (2003) On the shape of a set of points in the plane. *IEEE Trans Inf Theory* 29(4):551–559
- Guo Q, Quan Y, Jiang C (2019) Object pose estimation in accommodation space using an improved fruit fly optimization algorithm. *J Intell Rob Syst* 95(2):405–417
- Hinterstoisser S, Cagniart C, Ilic S et al (2012) Gradient response maps for real-time detection of textureless objects. *IEEE Trans Pattern Anal Mach Intell* 34(5):876–888
- Kehl W, Manhardt F, Tombari F et al (2017) SSD-6D: making RGB-based 3D detection and 6D pose estimation great again. In: IEEE international conference on computer vision
- Lou S, Jiang X, Scott PJ (2013) Application of the morphological alpha shape method to the extraction of topographical features from engineering surfaces. *Measurement* 46:1002–1008
- Pan WT (2012) A new Fruit Fly Optimization Algorithm: taking the financial distress model as an example. *Knowl-Based Syst* 26(2):69–74
- Santos RCD, Galo M, Carrilho AC (2019) Extraction of building roof boundaries from LiDAR data using an adaptive alpha-shape algorithm. *IEEE Geosci Remote Sens Lett* 16:1–5
- Shi J, Tomasi C (1994) Good features to track. In: IEEE conference on computer vision and pattern recognition, pp 593–600
- Wang Q, Chen X (2019) Building contours extraction from LIDAR data based on alpha shapes algorithm. *Intell Build Smart City* 2:23–25
- Wang C, Xu D, Zhu Y et al (2019) DenseFusion: 6D object pose estimation by iterative dense fusion. In: IEEE conference on computer vision and pattern recognition
- Xiang Y, Schmidt T, Narayanan V et al (2017) PoseCNN: a convolutional neural network for 6D object pose estimation in cluttered scenes. In: Robotics: science and systems
- Zhu C, Zhang X, Hu B et al (2008) Reconstruction of tree crown shape from scanned data. In: Proceedings of the 3rd international conference on technologies for e-learning and digital entertainment. Edutainment '08, pp 745–756. Springer-Verlag, Berlin

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.