Send Orders for Reprints to reprints@benthamscience.ae

The Open Automation and Control Systems Journal, 2015, *7*, 1875-1881

Recognizing Furniture Supported by Legs Based on Vision Sensor

Qing Ma[1,*], Chuanbo Zhou[2], Xin Wang[1] and Shilong Xu[3]

[1]Department of Control Science and Engineering, University of Shandong, Jinan, Shandong 250061, China; [2]Shandong Suntong Electrical Technology Co., Ltd, China; [3]Jinan Youbai Electronic Technology Co., Ltd, China

Abstract: A method of object recognition for objects with special topological structure is developed based on top-hat transformation for service robot. The objects with four legs are easily detected and recognized by service robot with laser ranger array by this method. First, the top-hat transformation in one dimension is reviewed, then the recognition strategy of self-adapting threshold for objects with special topological structure is proposed, and finally, the general data process for object recognition and position is proposed and analyzed. Experimental results show that the process of object recognition based on top-hat transformation proposed in this article is an effective and accurate application.

Keywords: Object recognition, self-adapting threshold, service robot, top-hat transformation.

1. INTRODUCTION

When robots carry on mission planning and execution, they not only need structural information about the environment for navigation and self-localization, but also need to understand semantic information of the environment to possess certain social skills. The semantic maps [1, 2] come up to capture the human point-of-view of robotic environments, which enable the high-level and more intelligent robot development and the human-robot interaction as well. To realize the above function, semantic map should include detailed information about the functional properties of rooms and objects [3-5]. For example, if a man needs a “chair”, the robot should find the semantic symbol and the corresponding location of the chair in its semantic map at first. Consequently, object recognition is the basis of advanced function for service robot to build the semantic map.

When the service robot mainly works in special environment such as wards in hospital, conference rooms, dining rooms or other indoor environment, many target objects such as beds, chairs, desks with four legs are faced. Recognition task for these items is very important for semantic map building. If the robot can recognize commands such as ‘this is a table’ and ‘that is a chair’, the robot can serve as being the user friendly.

Many techniques have been proposed and are now in use for object recognition; however, the recognition of the object with four legs is difficult and little reported. Computer vision is a common method for object recognition because image contains enough information and details for human to finish it, and many methods or algorithms have been proposed to solve the problem. But some problems still exist [6-9]. Firstly, in complex indoor environment, it’s hard to select candidate objects for various target objects. Secondly, it’s difficult to accurately extract shapes from objects because of different object postures and noises or illuminations from the background. Thirdly, Sample library should be built before recognition, for example, a database of 7000 pictures for 100 objects is already built in [8], which brings much work for developers. So, computer vision is not a good choice for service robot to finish recognition tasks because a large database needs to be achieved before. Besides computer vision for object recognition based on features of objects, there is another method proposed by Hao Wu, for service robot to recognize target object based on QR Code [9]. In this method, the information of object is recorded by QR code which is stacked on target object with artificial landmarks. It’s a reliable and effective method, but the object is recognized beforehand and all the information of the object is encoded to QR code by programmers. It requires more time for programmers to help the service robot finish the artificial recognition tasks.

Using laser sensor, a new method based on top-hat transformation and self-adapting threshold has been proposed for service robot to finish parts of object recognition tasks in our work. The experiments proved that all recognition tasks for furniture with four legs could be finished accurately and rapidly by service robot, which will save large amounts of time for developers.

2. OBJECT RECOGNITION BASED ON TOP-HAT TRANSFORMATION

2.1. Review of Top-Hat Transformation in One Dimension

Top-hat transformation is a morphological algorithm based on erosion operation and dilation operation. It is commonly used in medical imaging or agriculture processing and recognition, such as counting of blood cells, or detecting and counting of sample seeds etc. It is the high-pass filter in image processing that eliminates the signal with low frequency background, such as uneven illumination of sunlight or lamp. Top-hat transformation in one dimension is effective in weakening the background signal with low frequency and maintaining the effective signal of target objects. Let F and S represent the original signal from laser ranger array and a structuring element respectively. Erosion operation $F \oplus S(t)$ is defined below [10]:

*Address correspondence to this author at the Department of Control Science and Engineering, University of Shandong, Jinan, Shandong, 250061, P.R. China; Tel: +86053188393461; E-mail: maqing@sdu.edu.cn

1874-4443/15 2015 Bentham Open

$$F \oplus S(t) = \min\{F(x^0 - x') - S(x') \big| -x^0 < x' < x^0\} \quad (1)$$

Correspondingly, the dilation operation $F \oplus S(t)$ is defined below:

$$F\Theta S(t) = \max\{F(x^0 - x') + S(x') \big| -x^0 < x' < x^0\} \quad (2)$$

And the opening operation and closing operation F○S and F●S are defined below:

$$F \circ S = (F\Theta S) \oplus S \quad (3)$$

$$F \bullet S = (F \oplus S)\Theta S \quad (4)$$

Let H represent the data after top-hat transformation, and the top-hat transformation is defined as follows [10]:

$$H = F - (F \circ S) \quad (5)$$

In image processing, opening operation is used for brightening some regions and closing operation is used for darkening holes in images. The top-hat transformation in one dimension is used as a high-pass filter. The structuring element S is designed as a semicircle with the radius r. The signal with frequency lower than $1/r$ is weakened by filter, and all signals with frequency higher than that are maintained. The top-hat transformation is the fundamental filter for object recognition for service robot. Comparing previous method based on machine learning or object database, the proposed method can detect new objects accurately without previous learning or database building.

Let S be the structuring element of top-hat transformation, so each point S_i with position (x_i, y_i) of all the n points on S is on the circle with a special radius r and center (x_0, y_0) as shown in the following equation:

$$x_i = x_0 + r * \cos(\frac{i}{n} * 180) \quad (6)$$

$$y_i = y_0 + r * \sin(\frac{i}{n} * 180) \quad (7)$$

Fig. (**1**) shows a structuring element with radius 50.

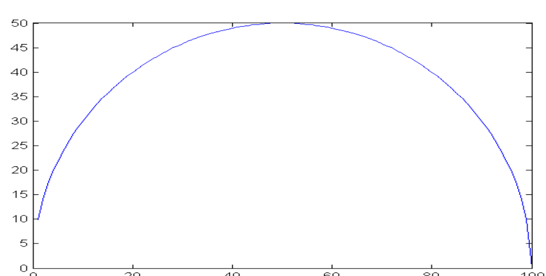

Fig. (1). A structuring element S.

2.2. Laser Ranger Array Processing of Service Robot

Laser ranger array is an important part of the multi-sensor system in service robot. In our service robot platform, we placed a laser ranger array in front of the body of robot. The array of laser ranger can send feedback with an array of length from the laser ranger array to the obstacle in polar coordinates. The angle of laser rays is from 0 to 180. There are 1080 laser rays in the laser ranger ray, and it can send feedback of a ray of 1080 focal length. The refresh frequency is about 7.69Hz. However, we could not use all of the data from the laser ranger array. So, we select from No. 280 to No. 800, which is ranged from angle 25 to 155 to avoid the influence of the body structure of the service robot.

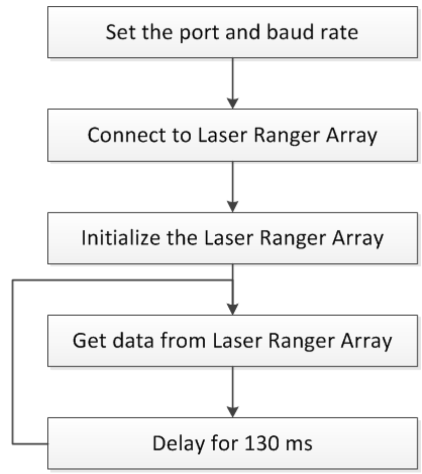

Fig. (2). Flow chart of laser ranger array.

2.3. General Processing Steps of Object Recognition

In the software platform, we also built a real-time module for the laser ranger array to return data array for different applications or tasks. The flow chart laser Fig. (**2**) shows how the laser ranger array collects the data and returns the values of lengths to the service robot. In the module of laser ranger array, the central computer of the service robot first chooses the communication port and sets the baud rate, then connects to the laser ranger array and starts initialization. Then, service robot will get the data from laser ranger array every 130 minutes. It's a real-time module and will return the data array with a frequency of 7Hz. The module of laser ranger array is the basis for not only object recognition task but also for many advanced functions or tasks such as navigation, obstacle avoidance and map building. If the sampling frequency is too high, the module of laser ranger array will occupy too much CPU and RAM. And, if the frequency is too low, the accuracy of other module will decrease. Let the laser ranger array sample a group of data from a chair, that is a sample of target object with four legs. Fig. (**3**) shows a group of data and Fig. (**4**) is a picture captured by camera on service robot to show the posture and position of the chair.

The general steps of object recognition method are shown below:

(1) Connect and initialize the laser ranger array to prepare for laser ranger array's task;

(2) Get data from laser ranger array and store the data in the RAM;

(3) Use smoothed filter to process the data and delete the noise signals. The smoothed filter is used to decrease the noise signals "1" and maintain the effective information. Let $f_{in}(i)$ as the data in laser ranger array with number i and $f_{out}(i)$ as the output of the smoothed filter, and the smoothed filter is represented below:

$$f_{out}(i) = \begin{cases} f_{in}(i) & \text{if } f_{in}(i) \neq 1 \\ f_{in}(i-1) & \text{if } f_{in}(i) = 1 \end{cases} \quad (8)$$

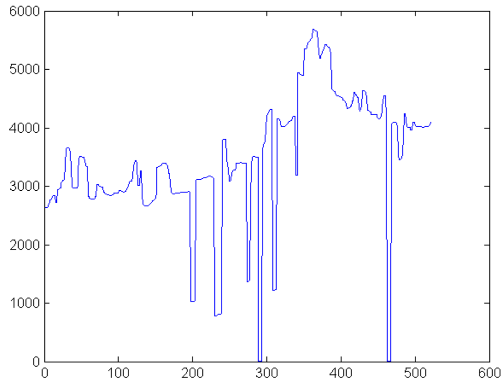

Fig. (3). Data of chair from laser ranger array.

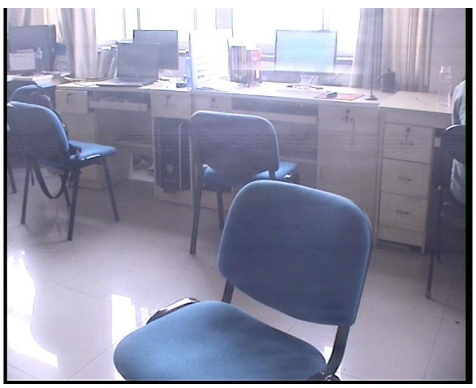

Fig. (4). Picture of target object.

(4) Find the maximum data, and then reverse data for top-hat transformation and build the structuring element for top-hat transformation, then use top-hat transformation to process the data to decrease the background signal and information. Let the *f*, *h* and *s* be the input and output of top-hat filter, as well as the structuring element of the top-hat transformation, respectively. The building of structuring element and top-hat transformation is shown in equations (5) to (7).

(5) Choose the threshold and binarize the data. Let the T denote the threshold of the binarization. The process of binarization is shown in equation (9).

$$t_{out}(i) = \begin{cases} 1 & \text{if } f_{out}(i) > T \\ 0 & \text{if } f_{out}(i) < T \end{cases} \quad (9)$$

(8) If the number of square wave is more than 3, then go to the next step, or ask for artificial recognition;

(9) Transfer the data from polar coordinate to rectangle coordinate for calculating and checking the results. Let the $f_{out}(i)$ denote the length from one point *i* of object to laser ranger, and $a_{out}(i)$ denote the angle of the point *i*. The transformation from polar coordinate to rectangle coordinate is shown in the equation (10) below:

$$\begin{cases} x(i) = f_{out}(i) * \cos[a_{out}(i)] \\ y(i) = f_{out}(i) * \sin[a_{out}(i)] \end{cases} \quad (10)$$

(10) The four points P_1,P_2,P_3,P_4 that make the lines A_1(by P_1,P_2), A_2(by P_2,P_4), A_3(by P_3,P_4), A_4(by P_1,P_3) are perpendicular to the adjacent lines, parallel with opposite line and the length of each line is equal to the opposite line; so that the four points create a rectangle. If not, ask for artificial recognition;

(11) Update the information of the object library and return to the current tasks.

The flow chart of the object recognition method is shown in Fig. (**5**).

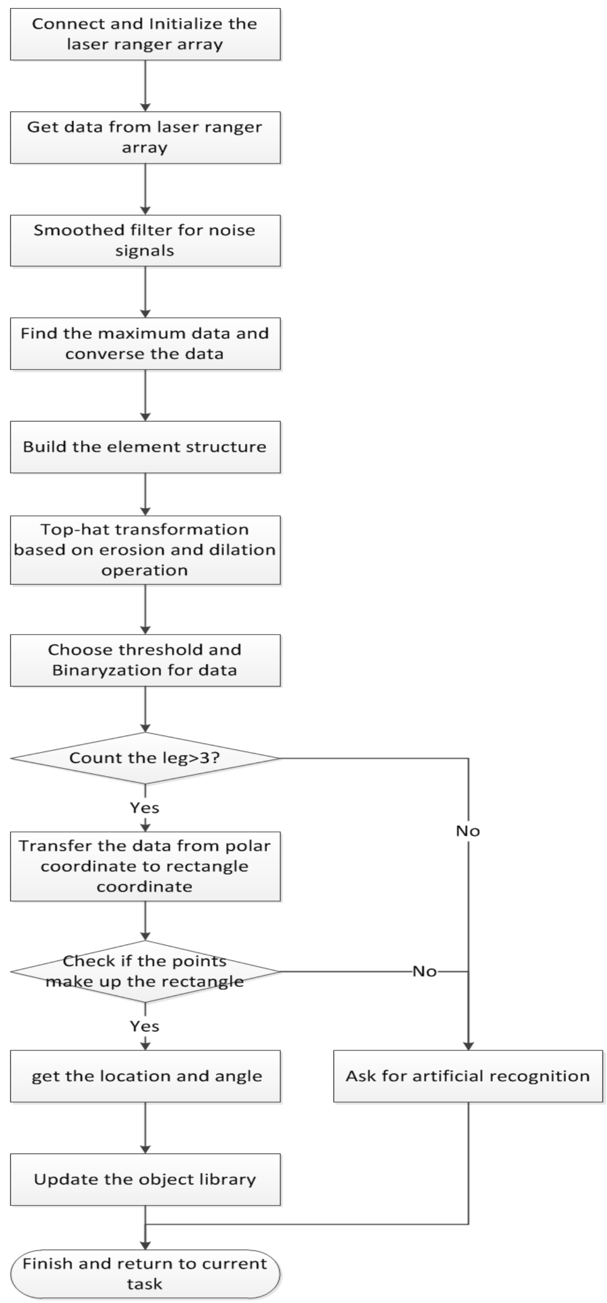

Fig. (5). Flow chart of object recognition.

In Fig. (**3**), some of noise signals '1' fall within the range from 280-300 and 450-470. Smoothed filter is used initially to remove the noise signal '1' from the data array. The effective signal is four groups of data with minimum values compared with complex background signal. So, the next step is to find the maximum of data array and reverse the data, and the task is switched to finding the maximum values from the background. Most valid information is from 5 to 20, so the structuring element *S* is designed as a semicircle with radius of 50 to ensure all the valid information is maintained. After the top-hat transformation, there are four obvious wave crests, each of which is one of the four legs of the chair and other background noises are decreased. Next, we will set the threshold for data binarizing as $0.5*\max\{l|l\in \text{mat}\}$. In fact, we set the threshold based on different rules. If the environment is large enough (larger than $20m^2$), then we set the threshold as the half of the maximum length $0.5*l$, and if the environment is smaller than $20m^2$, we set the threshold as (l-50)cm. After data analyzing, we can finally get the wave shape of data, and each square wave represents a group of data of each leg. We count the number of square waves, and if it is more than 3, the target object can be tentatively judged as a chair.

The next task is to calculate the length of each side of the chair, the posture of the chair and the center point of the chair with the data of square wave for building semantics map. First, we need to transfer the data from polar coordinate to rectangle coordinate, so that we get four groups of data, each of which is corresponding to each of the square wave. We can get the mean value (x,y) of each group of the data, which represents the position of four points P_1,P_2,P_3 and P_4. The four sides of rectangle made by the four points are A_1(by P_1,P_2), A_2(by P_2,P_4), A_3(by P_3,P_4) and A_4(by P_1,P_3). If each two gradients of two neighboring sides multiply to give about -1 and the lengths of each two opposite sides are equal, we can draw the conclusion that each two opposite sides are perpendicular to each other. At last we can get the length of each two legs of the chair by the Pythagorean Theorem, as well as the position and posture of target object. As for objects with four legs, the recognition strategy is set as follows: if each side of the object is less than 50cm, then it is recognized as a chair; if any two sides of the object are longer than 50cm but shorter than 120cm, then it is recognized as a desk; and if any two sides of the object are longer than 120cm, it is recognized as a bed. Special strategy for object recognition can be set by users in the interface of platform on service robot. If the type of object is with special topological structure, the graphic of the structure can be put into service robot for object recognition task based on top-hat transformation.

If the service robot works in a complex environment with too many similar target objects with special topological structure, after the top-hat transformation, we have found that there are too many square waves. To eliminate invalid information and maintain the suitable group of points, we introduced a method of self-adapting threshold. First, the rate of the threshold is set to the maximum length from 2 to lower by step of 0.05 until the number of square waves is decreased to 7 or less. Then, each 4 points of the 7 points are evaluated if they are the points of rectangle. If more than one rectangle is detected, then the rate is decreased until there is only one rectangle detected. If no rectangle is detected, then we may ask assistance of artificial recognition for help.

3. EXPERIMENTS

Fig. (**6**) is the result of signal processing after smoothed filter. Fig. (**7**) is the data after reversing. Fig. (**8**) is the structuring element *S* and Fig. (**9**) is the data after top-hat transformation. Fig. (**10**) is the data after binarization. The four square waves in Fig. (**10**) represents four groups of points of chair legs. Fig. (**11**) shows the data that is transferred from polar coordinate into rectangle coordinate.

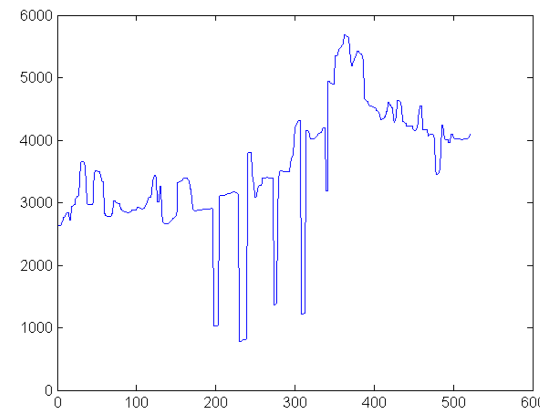

Fig. (6). Data after smoothing filter.

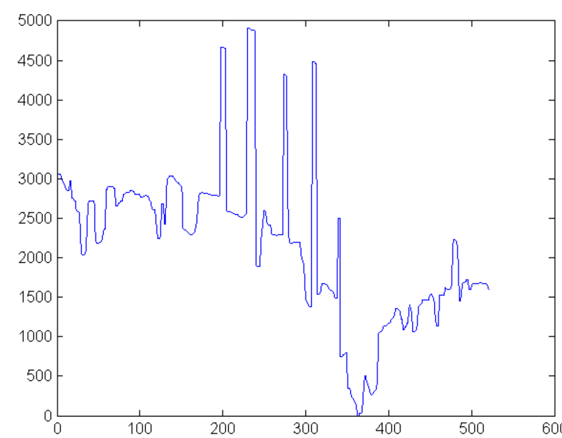

Fig. (7). Data after reversing.

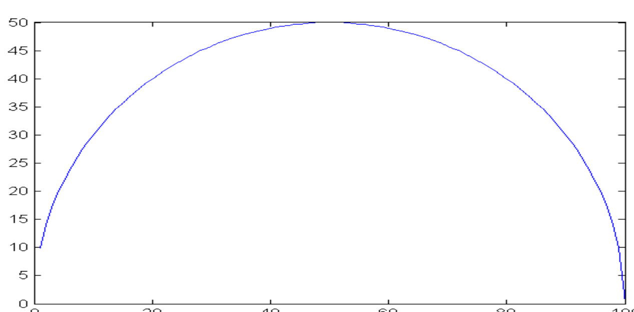

Fig. (8). Structuring element S.

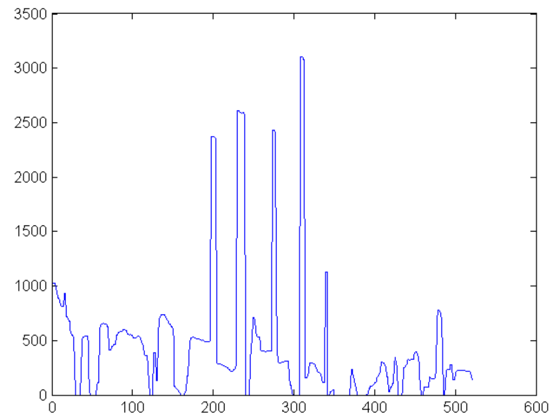

Fig. (9). Data after top-hat transformation.

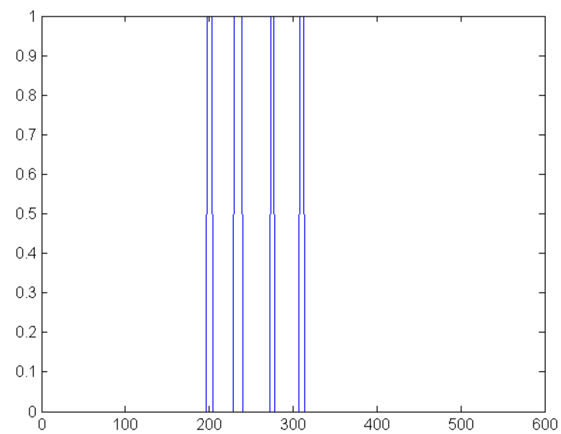

Fig. (10). Data after binarizing.

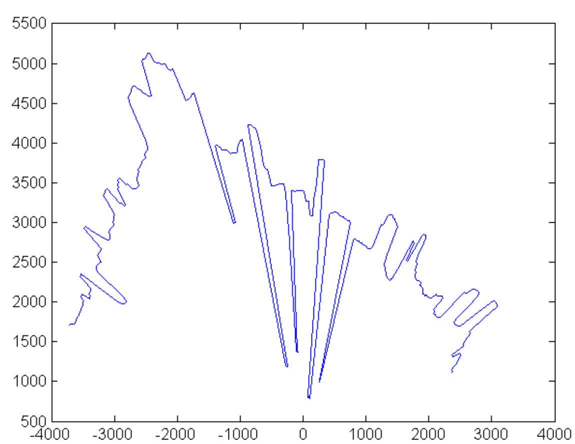

Fig. (11). Data transferred to rectangle coordinate.

The center position of four groups of data is shown in Table **1**.

Then, we can get gradients of each side of the chair edge:

$k_1(1,2)$=1.1113

$k_2(2,4)$=-1.1162

$k_3(4,3)$=1.0307

$k_4(3,1)$=-1.0764

And we verify that if each two gradients of two neighboring side are perpendicular:

k_1*k_2=-1.2405

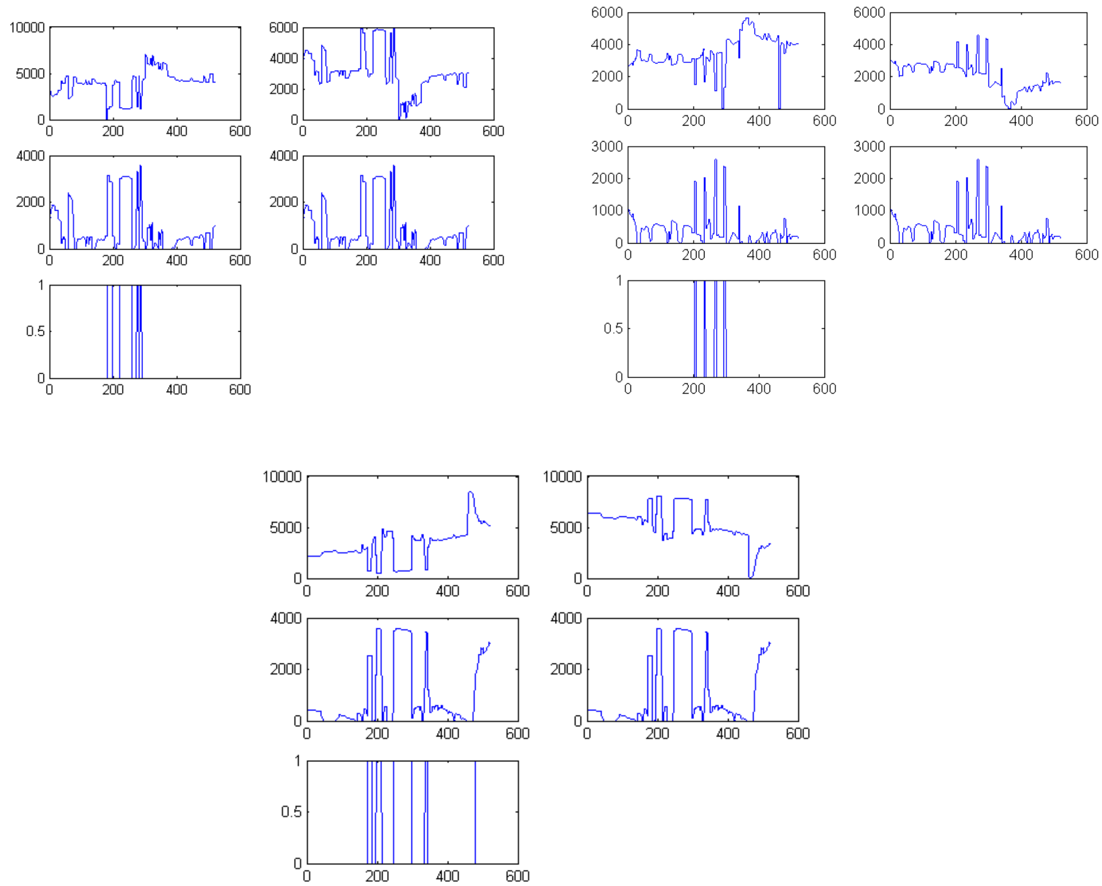

Fig. (12). Other samples of object recognition.

k_2*k_3=-1.1505

k_3*k_4=-1.1094

k_4*k_1=-1.1962

Each two gradients of two neighboring sides multiply that is about -1, which means each two sides are perpendicular. Then, we can get the distance of each two legs of the chair by the Pythagorean Theorem in millimeter:

L_1=264.7410

L_2=533.6242

L_3=254.3703

L_4=569.6627

After taking measurement of the chair, suitable data is obtained. Next, we can obtain the position and the posture angle of the chair, which are approximately accordant with the real value:

x=-0.7073

y=1089.2

a=6.0353

Fig. (**12**) shows other groups of data and results of each steps during data processing. All the target objects are recognized accurately.

Table 1. Center position of four groups points.

Group ID	Data ID	Mean X	Mean Y
1	197-203	265.8579	992.3953
2	230-239	88.7721	795.5999
3	273-277	-90.1656	1375.6129
4	308-313	-267.2935	1193.1002

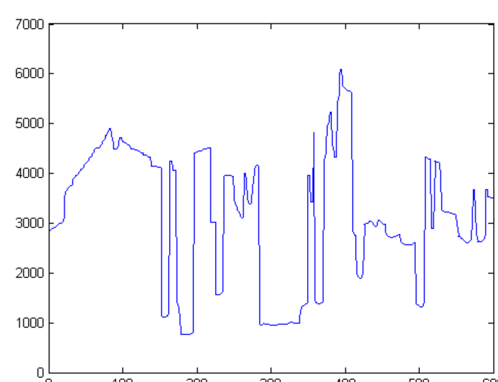

Fig. (13). Data of bed from laser ranger array.

Fig. (14). Picture of bed in the environment.

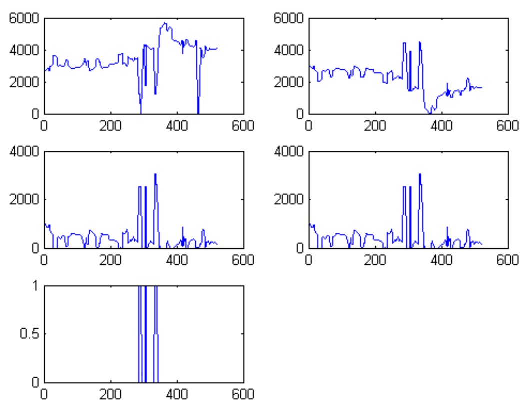

Fig. (15). The experimental result of object recognition for bed.

Another group of data for bed from laser ranger array is shown in Fig. (**13**). The picture of target object is captured by camera, which is shown in Fig. (**14**). The result is shown in Fig. (**15**). From the experimental result, the algorithm of top-hat transformation is found effective in different kinds of object recognition in complex environments.

CONCLUSION

The top-hat transformation is an effective method for service robot in recognizing objects. The valid signal is maintained for further recognition and the background signal is eliminated. Objects with special topological structure are accurately detected and have a fixed position in this method. The method proposed in this article is efficient and precise for recognition tasks of service robot. The amount of calculation and the complexity of the method is low enough that doesn't interfere other tasks or modules of service robot. Though limited types of objects can be detected, other types of objects can be recognized by transforming the method.

An important area for further research is 3D modeling based on laser ranger array or 3D cameras. 3D model is a detailed data with features that can be recognized with advanced algorithm in computer vision or signal processing. More types of objects would be recognized with more features in shape, topological structure, size etc. 3D modeling and data analyzing requires more computing capabilities and RAM of service robot, especially for dynamic objects, so algorithm optimization would become one of the central topics in object recognition.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

This work was financially supported by National Nature Science Foundation under Grant 61203330, and Shandong Nature Science Foundation under Grant ZR2012FM031, China Postdoctoral Science Foundation funded project 2013M540546, Shandong Postdoctoral Innovation Foundation 201203058.

REFERENCES

[1] A. Nüchter, and J. Hertzberg, "Towards semantic maps for mobile robots," *Robotics and Autonomous System*, vol. 56, no. 11, pp. 915-926, 2008.

[2] D.F. Wolf, and G.S. Sukhatme, "Semantic mapping using mobile robots, "*IEEE Transactions on Robotics*, vol. 24, no. 2, pp. 245-258, 2008.

[3] J.G. Rogers, and H. I. Christensen, "Robot planning with a semantic map," In: *IEEE International Conference on Robotics and Automation (ICRA)*, 2013, pp. 2239-2244.

[4] T. Wang, and Q. Chen , "Object semantic map representation for indoor mobile robots, In: *Proceedings of the 2011 International Conference on System Science and Engineering*, 2011, pp. 309-313.

[5] D.W. Ko, C. Yi, and I.H. Suh, "Semantic mapping and navigation: A Bayesian approach," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2013, pp. 2630-2636.

[6] D.G. Lowe, "Object recognition from local scale-invariant features," Computer vision, *Proceedings of the 7th IEEE International Conference on Robotics*, vol. 2, pp. 134-140, 1999.

[7] S. Belongie, M. Jitendra, and P. Jan, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509-522, 2002.

[8] C. Nakajima, N. Itoh, M. Pontil, and T. Poggio, "Object recognition and detection by a combination of support vector machine and rotation invariant phase only correlation," In: *15th International Conference on Pattern Recognition*, 2000, pp. 787-790.

[9] H. Wu, G.H. Tian, Y.H. Xue, and T.T. Zhang, "QR code based semantic map building in domestic semi-unknown environment," *Pattern Recognition and Artificial Intelligence,* vol. 23, no. 4, pp. 464-470, 2010.

[10] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, New York, 1982.

Received: May 26, 2015 Revised: July 14, 2015 Accepted: August 10, 2015