International Journal of Control and Automation
Vol. 7, No. 12 (2014), pp. 311-322
http://dx.doi.org/10.14257/ijca.2014.7.12.29

Four-Leg Object Recognition for Service Robot Based on Top-hat Transformation

Hao Wu, Guohui Tian, Xinran Wang and Fei Lu

School of Control Science and Engineering, Shandong University
wh911@sdu.edu.cn

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 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: *Top-hat transformation, self-adapting threshold, service robot, object recognition*

1. Introduction

When robots carry on mission planning and execution, they not only need structural information about the environment to 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 robot environments, which enable the high-level and more intelligent robot development and the human-robot interaction as well. To realize above function, semantic map should include detailed information about the function property 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 this is a table and that is a chair, the robot can serve the user friendly.

Many techniques are proposed and 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 that for human to finish it, and many methods or algorithms are proposed to solve the problem. But some problems still exist [7-9]. Firstly, in complex indoor environment, it's hard to select candidate objects for various target objects. Secondly, it's difficult to extract shapes from objects accurately because of different object postures and noises or illuminations from background. Thirdly, Sample library should be built before recognition, for example, a database of 7000 pictures for 100 objects is

ISSN: 2005-4297 IJCA

built in [8], and it 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, another method for service robot to recognize target object based on QR Code is proposed by Hao Wu [9]. In her method, the information of object is recorded by QR code which is sticked on target object with artificial landmarks. It's a reliable and effective method, but the object is recognized beforehand and all the information of 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 is 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, it will save large amount 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's common in medical image or agriculture processing and recognition, such as counting for cells in blood, or detecting and counting for sample seeds. It's a high-pass filter in image processing that eliminates the signal with low frequency in 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. Erosion operation $F \oplus S(t)$ are defined below [10]:

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

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

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

And the opening operation and closing operation $F \circ S$ and $F \bullet S$ are defined below:

$$F \circ S = (F \Theta S) \oplus S \tag{3}$$

$$F \bullet S = (F \oplus S) \Theta S \tag{4}$$

Let H represents the data after top-hat transformation, and the top-hat transformation is defined by [10]:

$$H = F - (F \circ S) \tag{5}$$

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 radius of r. The signal with frequency lower than $1/r$ is weaken by filter, and all signals with frequency higher than that is 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 as the structuring element of top-hat transformation, so each point Si with position (x_i, y_i) of all the n points on S is on the circle with special radius r and center (x_0, y_0) in following equation:

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

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

Figure 1 shows a structuring element with radius 50.

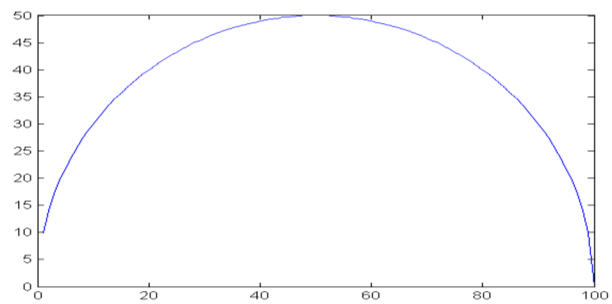

Figure 1. A Structuring Element *S*

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 place a laser ranger array in the front of the body of robot. The arrays of laser ranger can 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 feedback a ray of 1080 lengths. 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.

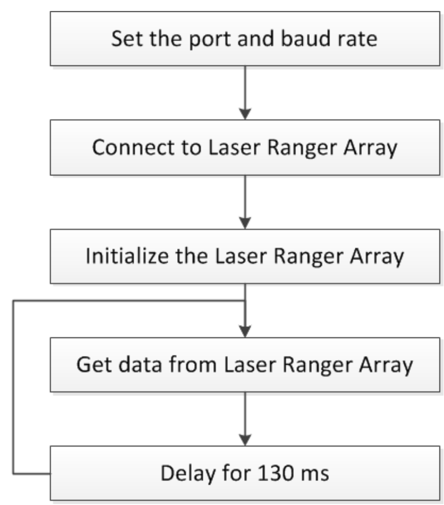

Figure 2. Flow Chart of Laser Ranger Array

2.3 General Processing Steps of Object Recognition

In the software platform, we also build a real-time module for the laser ranger array to return data array for different application or other tasks. The flow chart laser Figure 2 shows how the laser ranger array collects the data and return 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 ms. It’s a real-time module and will return the data array in frequency of 7Hz. The module of laser ranger array is the basic for not only object recognition task but also 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 samples a group of data from a chair, a sample of target object with four legs. Figure 3 shows a group of data and Figure 4 is a picture captured by camera on service robot to show the posture and position of the chair.

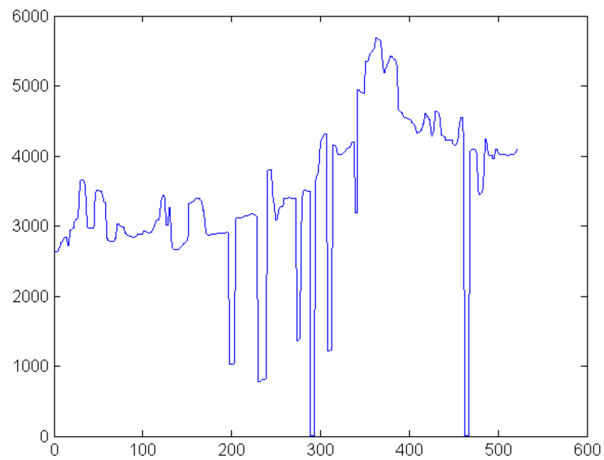

Figure 3. Data of Chair from Laser Ranger Array

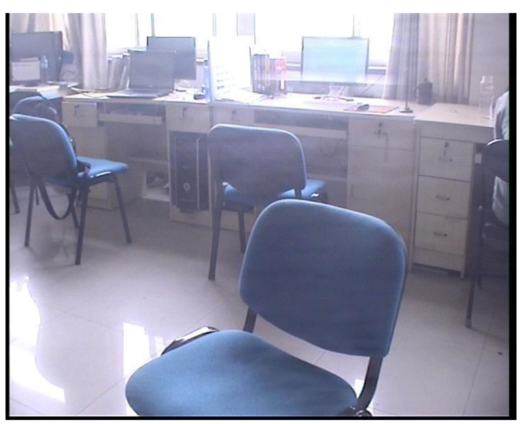

Figure 4. Picture of Target Object

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:

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

(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* the input and output of top-hat filter, as well as the structuring element of the top-hat transformation. The building of structuring element and top hat transformation is shown in equation (5) to (7).

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

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

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

(7)Transfer the data from polar coordinate to rectangle coordinate for calculation and checking for 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 equation (10) below.

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

(8)The four points P_1,P_2,P_3,P_4 make up the line 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 with the adjacent lines, parallel with opposite line and the length of each line is equal to the opposite line, then the four points make up a rectangle. If not, ask for artificial recognition;

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

The flow chart of the object recognition method is shown in Figure 5.

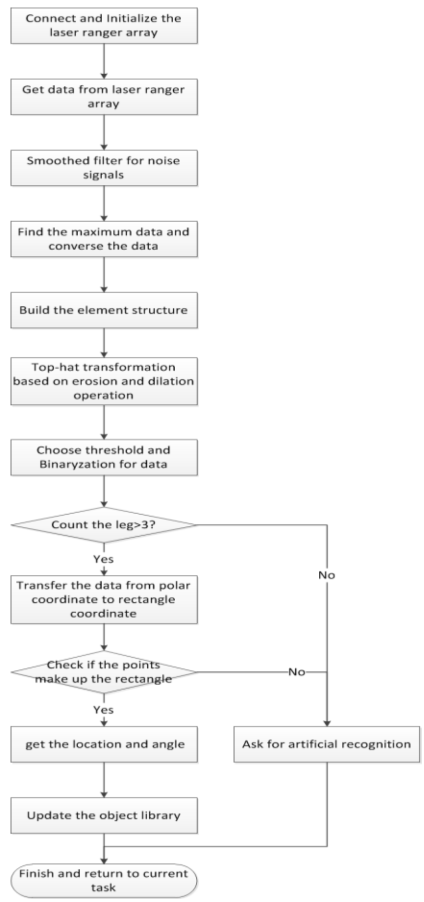

Figure 5. Flow Chart of Object Recognition

In Figure 3, some of noise signals '1' are shown in the range from 280-300 and 450-470. Smoothing filter is used first to remove the noise signal '1' from the data array. The effective signal is four groups of data with minimum values comparing 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 by 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 for the number of square wave, 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, and 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 up 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 neighbor sides multiply is about -1 and the lengths of each two opposite sides are equal, we can get the conclusion that each two opposite sides are perpendicular. 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 that if each sides of the object is less than 50cm, then it's recognized as a chair; if any two sides of the object is longer than 50cm but shorter than 120cm, then it's recognized as a desk; if any two sides of the object is longer than 120cm, it's recognized as a bed. Special strategy for object recognition can be set by users in the interface of platform on service robot. If a type of object with special topological structure, the graphic of the structure can be input 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, and after the top-hat transformation, we've found that there are too many square waves. To eliminate invalid information and maintain the suitable group of points, we introduce a method of self-adapting threshold. First set the rate of the threshold 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 we check each 4 points of the 7 points that if they are the points of rectangle. If more than one rectangle is detected, then we decrease the rate until there is only one rectangle is detected. If no rectangle is detected, then we may ask assistance of artificial recognition for help.

3. Experiments

Figure 6 is the results of signal processing after smoothing filter, Figure 7 is the data after reversing. Figure 7 is the structuring element *S* and Figure 8 is the data after top-hat transformation. Figure 10 is the data after binarize. The four square waves in Figure 9 represents four groups of points of chair legs. Figure 11 is the data that transferred from polar coordinate into rectangle coordinate.

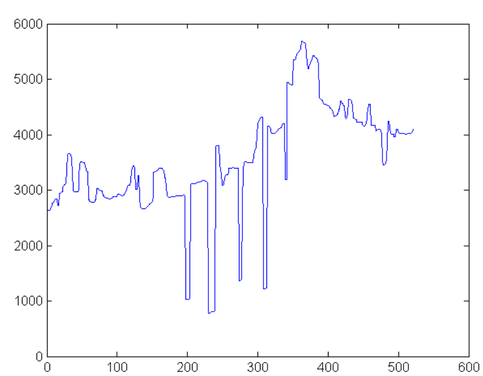

Figure 6. Data after Smoothing Filter

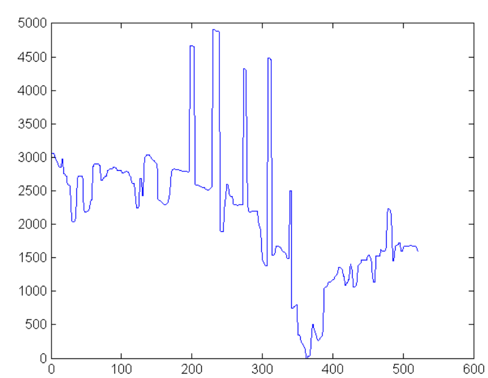

Figure 7. Data after Reversing

International Journal of Control and Automation
Vol. 7, No. 12 (2014), pp. 311-322
http://dx.doi.org/10.14257/ijca.2014.7.12.29

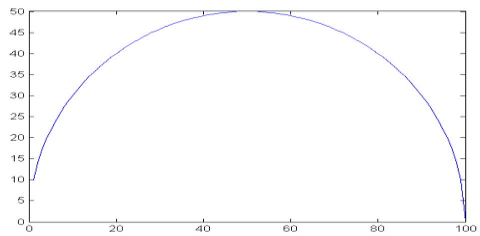

Figure 8. Structuring Element S

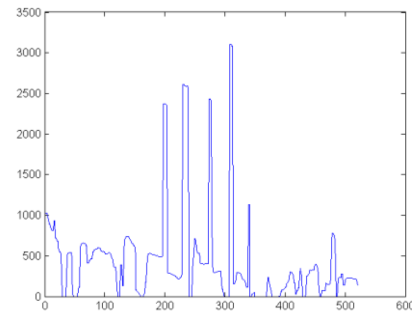

Figure 9. Data after Top-hat Transformation

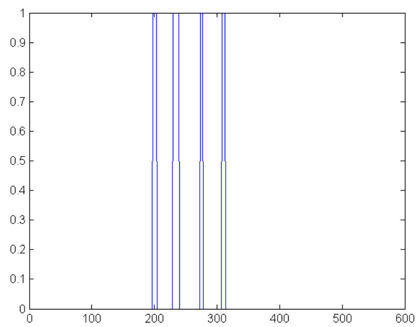

Figure 10. Data after Binarizing

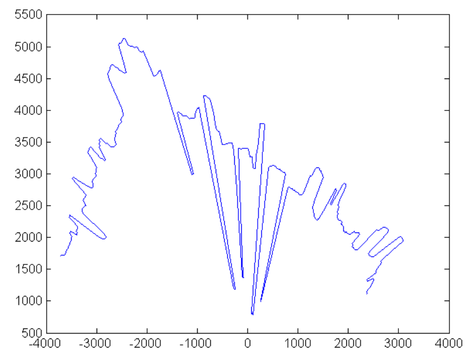

Figure 11. Data Transferred to Rectangle Coordinate

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

Chart 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

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 neighbor side is perpendicular:

$$k_1*k_2=-1.2405$$
$$k_2*k_3=-1.1505$$

$$k_3*k_4=-1.1094$$
$$k_4*k_1=-1.1962$$

Each two gradients of two neighbor sides multiply is about -1, that 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 our measurement to the chair, the data is suitable. Next we can obtain the position and the posture angle of the chair, which are approximately accordant toward the real value:

$$x=-0.7073$$
$$y=1089.2$$
$$a=6.0353$$

Figure 12 shows other groups of data and results of each steps during data processing. All of the target object is recognized accurately.

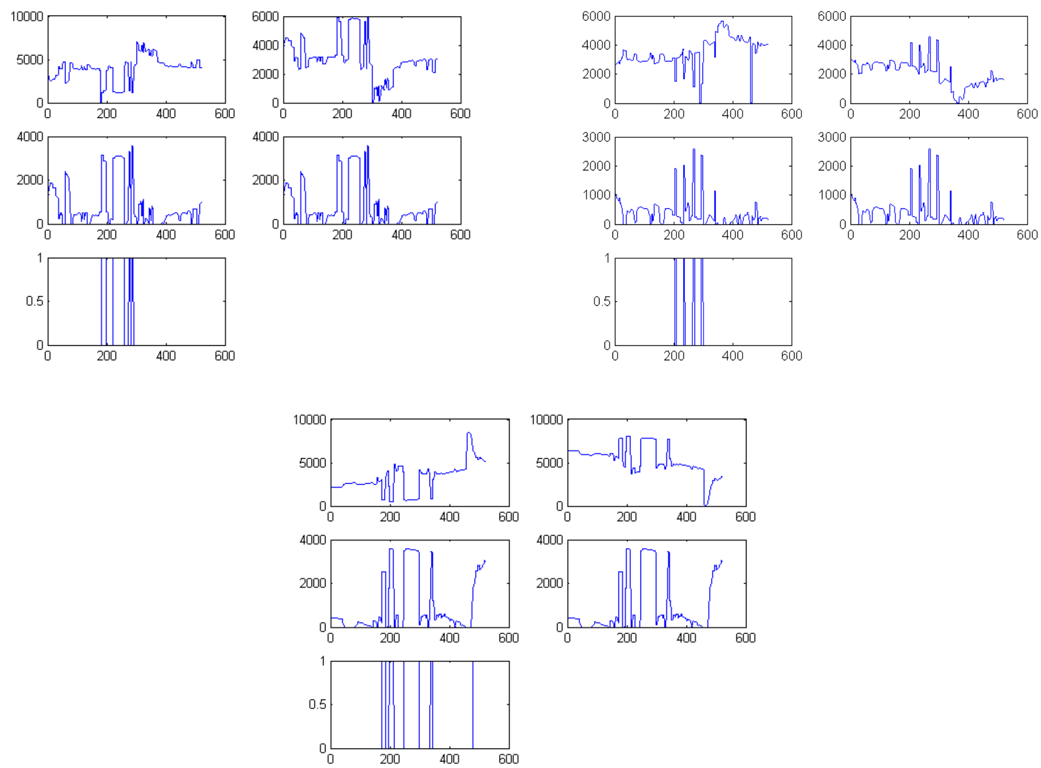

Figure 12. Other Samples of Object Recognition

Another group of data of bed from laser ranger array is shown in Figure13. The picture of target object is captured by camera in Figure 14. The result is shown in Figure 15. From the experimental result, the algorithm of top-hat transformation is effective in different kinds of object in complex environment.

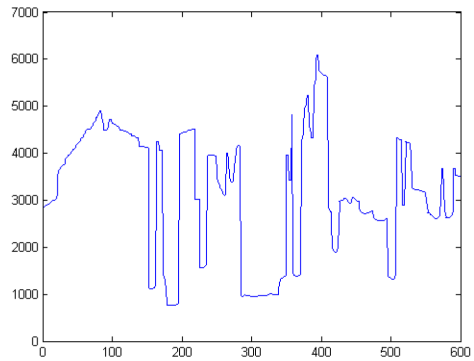

Figure 13. Data of Bed from Laser

Figure 14. Picture of Bed in the Ranger Array Environment

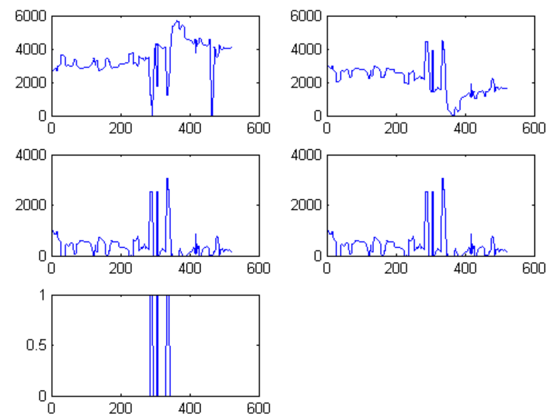

Figure 15. The Experimental Result of Object Recognition for Bed

4. Conclusion

The top-hat transformation is an effective method for service robot in recognition. The valid signal is maintained for further recognition and the background signal is eliminated. Objects with special topological structure is accurately detected and 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 object 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 object would be recognized with more features in shape, topological structure, size or else. 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.

Acknowledgements

Research supported by National Nature Science Foundation under Grant 61305122, 61203330, and Shandong Nature Science Foundation under Grant ZR2012FM031, ZR2011FM011 and ZR2010FM007, China Postdoctoral Science Foundation funded project 2013M540546, Independent Innovation Foundation of Shan dong university 2012TS078, Shandong Postdoctoral Innovation Foundation 201203058.

Reference

[1]. A. Nüchter and J. Hertzberg, “Towards semantic maps for mobile robots”, Robotics and Autonomous System, vol. 56, (**2008**) pp. 915-926.

[2]. D. F. Wolf and G. S. Sukhatme, “Semantic Mapping Using Mobile Robots”, IEEE Transactions on robotics, vol. 24, (**2008**), pp. 245-258.

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

[4]. W. Tingqi and C. Qijun, “Object semantic map representation for indoor mobile robots”, In proceedings of the 2011 International Conference on System Science and Engineering, ICSSE , (**2008**), pp. 309-313.

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

[6]. D. G. Lowe, "Object recognition from local scale-invariant features.", The proceedings of the seventh IEEE international conference on Computer vision, vol. 2, (**1999**).

[7]. S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts." IEEE Transactions on Pattern Analysis and Machine Intelligence, (**2002**), pp. 509-522.

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

[9]. H. Wu, *et al.*, "QR Code Based Semantic Map Building in Domestic Semi-Unknown Environment." Pattern Recognition and Artificial Intelligence, vol. 23, (**2008**), pp. 464-470.

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