3D Semantic Map-Based Shared Control for Smart Wheelchair

Zhixuan Wei[1,2], Weidong Chen[1,2], and Jingchuan Wang[1,2]

[1] Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China
[2] State Key Laboratory of Robotics and System (HIT), Harbin 150001, China
{zhixuan.wei,wdchen,jchwang}@sjtu.edu.cn

Abstract. The previous perception and control system of smart wheelchairs normally doesn't distinguish different objects and treats all objects as obstacles. Consequently it is hard to realize the object related navigation tasks such as furniture docking or door passage with interference from the obstacle avoidance behavior. In this article, a local 3D semantic map is built online using a low-cost RGB-D camera, which provides the semantic and geometrical data of the recognized objects to the shared control modules for user intention estimation, target selection, motion control, as well as parameters adjusting of weight optimization for addressing different target. With the object information provided by 3D semantic map, our control system can choose different behaviors according to user intention to implement object related navigation. A smart wheelchair prototype equipped with a Kinect is developed and tested in real environment. The experiments showed that the 3D semantic map-based shared control can effectively enhance the smart wheelchair's mobility, and improve the collaboration between the user and the smart wheelchair.

Keywords: Smart Wheelchair, 3D Semantic Map, Shared Control.

1 Introduction

To improve the mobility of the smart wheelchair and the collaboration between the user and the smart wheelchair is a currently important research topic worldwide, especially facing unknown indoor environment and accurate tasks. Smart wheelchair is required to cognize the environment, to estimate the intention of user, and to timely adjust the control strategies, so as to achieve accurate and complex operations such as door passage and furniture docking. Previous shared control cannot solve these problems, because it has weak environment perception, which means that it cannot distinguish different objects so that treats all objects as obstacles. Taking door passage as an example, in order to ensure safety, the best method is to pass through it along the perpendicular bisector of the door, but the previous control algorithm does not guarantee this. Another example is the docking into the table, which need to detect the table and determine the docking position and orientation. Previous shared control is almost impossible to solve such problem. In this article, we used shared control and 3D

C.-Y. Su, S. Rakheja, H. Liu (Eds.): ICIRA 2012, Part II, LNAI 7507, pp. 41–51, 2012.

semantic map on smart wheelchair to improve its environment perception and hence mobility and collaboration.

Various shared control has been proposed since 1990s [1]. They can be divided into two categories according to the control level a user takes part in: behavior level sharing and planning level sharing.

Behavior level sharing is a commonly used method in the early stage. There are usually two ways for a wheelchair to cooperate with a human. The first way is that a wheelchair goes towards a direction that a user points out and the assistive system provides some obstacle avoidance algorithm to ensure safety [2]. In the second way, the user's commands are treated as a behavior which is executed with other autonomous behaviors (e.g. obstacle avoidance behavior, wall follow behavior).

Planning level sharing takes the user's intention into account while doing planning. The wheelchair follows orders coming from a planner, and user expresses his or her intention by moving the joystick. When the user's intention conflicts with the planner's order, the control system will modify the user's command [3] or re-plan the task [4], [5]. The user's intention of doing a certain task (e.g. door passage) is measured by defining intention prediction functions [4].

A new kind of shared control method was recently proposed in [6] and [7]. They defined an efficiency function to evaluate the user's control ability and adjusted the user's control weight according to the function value. Inspired by above works, in our previous work [8], we proposed a minimax algorithm for optimizing the weights of both commands of user and machine. All methods mentioned above, however, don't distinguish objects in environment but consider them all as obstacles.

Many 3D technologies have been applied to smart wheelchair since 2005. Stereo vision-based SLAM is used for the smart wheelchair navigation in [9]. But the maps only contained geometry information without object information. The 3D model is segmented into distinct potentially traversable ground regions and fitted planes to the regions in [10]. The planes and segments were analyzed to identify safe and unsafe regions and the information was captured in an annotated 2D grid map called a local safety map. But they still cannot distinguish different objects either.

Rusu et al. [11] proposed a novel framework for semantic 3D object model acquired from point cloud data. The functionality of this framework included robust alignment and integration mechanisms for partial data views, fast segmentation into regions based on local surface characteristics, and reliable object detection, categorization, and reconstruction. The computed models were semantic, i.e. they inferred structures in the data, which are meaningful with respect to the robot task. Such objects include doors, handles, supporting planes, cupboards, walls, and movable smaller objects. The point clouds are resulting from a 3D laser scanner. For smart wheelchair application, the mapping approach is still facing the issues on real-time computation, low-cost sensor and human-wheelchair cooperation.

In this article, system architecture of shared control for smart wheelchair is presented. A local 3D semantic map is online built with use of a low-cost RGB-D camera, which provides the semantic and geometrical data of the recognized objects to the shared control modules for user intention estimation, target selection, motion control, as well as parameters resetting of weight optimization for addressing different target, A smart wheelchair equipped with a Kinect is developed as experimental platform for studying the effectiveness of the proposed method.

2 System Architecture

As shown in Fig. 1, our approach is based on the previous shared control (below the dash). 3D semantic map contain the target information from 3D object detection and object feature extraction. At the same time, user intention is estimated to determine whether the user would like to reach the target. If not, the shared control will work as usual; if yes, the 3D semantic map will plan the motion to drive the wheelchair to the target, and the output of motion control (linear and angular velocity commands) will replace the output of joystick, meanwhile the 3D semantic map will adjust the internal parameters of the shared control to adapt to the different situations.

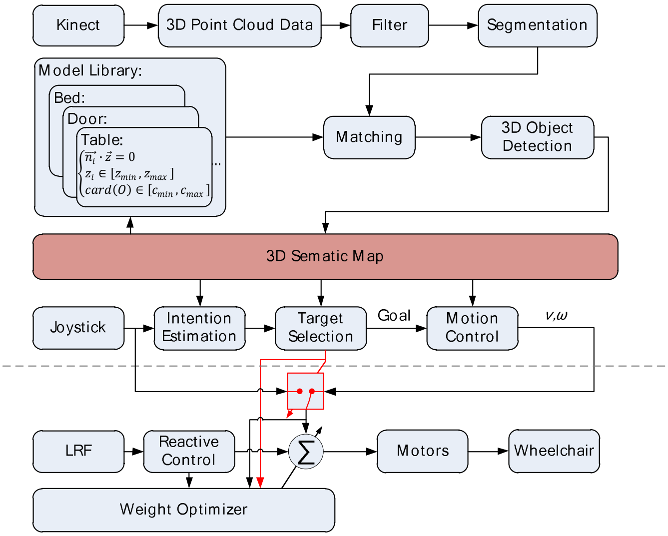


Fig. 1. System architecture

3 3D Semantic Map Building

In this article, we use shape-based method to build 3D semantic map. The point cloud data obtained from the RGB-D camera is firstly filtered and down-sampled to reduce the amount of data. Secondly, RANSAC algorithm is used to segment the data in accordance with the horizontal plane and vertical plane, then European clustering is used to make segmentation region finer. Finally, each region is matched using a priori model library in order to identify object, and to extract object feature for navigation.

Pass Through Filter. Pass through filter is used to reduce the amount of date by removing the useless points, such as the points which are too far or the points higher than the wheelchair and the user. O_i is the point cloud and p (x, y, z) is a point:

$$O_i = \{p(x, y, z) | x \in (x_1, x_2), y \in (y_1, y_2), z \in (z_1, z_2)\} \ . \tag{1}$$

Sparse Outlier Removal. Using RGB-D camera, measurement errors lead to sparse outliers which corrupt the results even more. They complicate the estimation of local point cloud characteristics such as surface normals or curvature changes, leading to erroneous values. The sparse outlier removal module corrects these irregularities by computing the mean μ and standard deviation σ of the nearest neighbor distances, and trimming the points which fall outside the $\mu \pm \alpha \cdot \sigma$ [11]. The value of α depends on the size of the analyzed neighborhood.

Down Sample. The point cloud data obtained from RGB-D camera has high resolution and uneven density, which increase the amount of data. Space can be divided into voxel grids with some scale, and each grid contains at most one point. Such treatment can reduce the resolution and uniform the point cloud.

Point Cloud Segmentation. In our approach, objects we interested in are usually structured by a set of planes, particular the planes perpendicular or parallel to the ground. We use RANCAS algorithm first to extract the above-mentioned planes [12].

RANSAC is an abbreviation for "RANdom SAmple Consensus". It is an iterative method to estimate parameters of a mathematical model from a set of observed data which contains outliers.

The planes perpendicular and parallel to the horizontal plane can be constrained by the following equation, where $\vec{n_i}$ is the normal of each point in the point cloud data, and z-axis perpendicular to the ground:

$$\vec{n_i} \times \vec{z} = 0 \ . \tag{2}$$

$$\vec{n_i} \cdot \vec{z} = 0 \ . \tag{3}$$

RANSAC doesn't consider the continuity of data. The extracted results usually contain many separate areas. Therefore, the Euclidean cluster is needed to get the biggest continuity region of the result of RANSAC.

Model Matching. Common objects of the indoor environment can usually be described with some common-sense constraints, which include the plane normal, the range of area and height, etc. For example, a table can be described as follow:

$$\begin{cases} \vec{n_i} \cdot \vec{z} = 0 \\ z_i \in [z_{min}, z_{max}] \\ card(O) \in [c_{min}, c_{max}] \end{cases} \ . \tag{4}$$

where $card(O)$ is the number of points in the plane. Because each voxel grid contains at most one point after down sample, the number of points can be estimated to a plane area. So $card(O) \in [c_{min}, c_{max}]$ is the constraint of the desktop area.

In last step, RANSAC extracts the plane $ax + by + cz = d$. Hence $x = 0$ and $y = 0$ will get $z_i = d/c$. And $\vec{n_i} \cdot \vec{z} = 0$ is the constraint using RANSAC to extract planes parallel to the horizontal plane.

As result the algorithm have detected the object and marked the points in the point cloud which belong to the object to build 3D semantic map. These objects are candidates for the target selection module.

Object Feature Extraction for Navigation. Since our aim is to improve the wheelchair mobility, the details should be extracted according to different objects. One of the important details is to obtain the current goal for motion control in navigation, such as the orientation and the midpoint of the door, the position and orientation for docking, and the size of free space of the table.

4 Shared Control

The share control has two key parts: the reactive control and the weight optimization. The reactive control provides basic obstacle avoidance using MVFH&VFF methods [3], [4]. The weight optimal algorithm optimizes three indicators which will be discussed in the following section to obtain weight of reactive control and user.

Weight Optimization. In our previous work [8], indicators of wheelchair's performance were proposed: *safety*, *comfort* and *obedience*. *safety* measures the probability of collision. *comfort* measures the variation of angular velocity. *obedience* measures the degree of obedience to the user's control intention. These indicators are defined as:

$$safety = 1 - \exp(-\alpha \cdot dis) \ . \tag{5}$$

$$comfort = \exp(-\beta|\omega - \omega_0|) \ . \tag{6}$$

$$obedience = \exp(-\gamma|\xi - \xi^*|) \ . \tag{7}$$

where, *dis* measured the distance between the wheelchair and the nearest obstacle in its path; ω and ω_0 are the desired and current angular velocity; ξ^* is the orientation of user command calculated from the user's input v_{mach} and ω_{mach}; ξ is the orientation of final command determined by v and ω; α, β and γ are constants.

The aim of weight optimization is to maximize all three indicators. However, these indicators are usually contradictory to each other. Therefore, there is no absolute optimum solution for maximize the three indicators at the same time. So we proposed of solving this problem is: always improve the smallest indicator among the three. In accordance with this principle we choose the minimax method to simplify this multi-objective optimization problem to a single objective one (Eq. 8).

$$\begin{cases} \max_\kappa(\min(safety, comfort, obedience)) \\ s.t. \\ v(t) = v_{user}(t) \\ \omega(t) = \kappa\omega_{user}(t) + (1-\kappa)\omega_{mach}(t) \\ 1 \geq \kappa \geq 0 \end{cases} . \tag{8}$$

where, κ and $(1-\kappa)$ is the user weight and the reactive control weight; $v(t)$ and $\omega(t)$ is the linear and angular velocity to be sent to the wheelchair. This equation means that finding the user weight is equivalent to finding the proper κ to maximize the objective function $\min(safety, comfort, obedience)$ under the restrictions stated after s.t.. As the linear velocity in MVFH is equal to $v_{user}(t)$ as long as there is no possible collision, we restrict $v(t)$ to be equal to $v_{user}(t)$.

Eq. 8 as a linearly constrained nonlinear programming problem, there is generally no analytical solution, since we use one-dimensional search algorithm to solve the optimization: First, use rough search algorithm to determine the interval that contain the maximum of the objective function $\min(safety, comfort, obedience)$; Second, implement Golden section search algorithm in the interval mentioned above to find the κ at maximum of the objective function.

Intention Estimation and Target Selection. Human should always be dominant in shared control. The command of machine plays a role of optimizing or revising user's order, which is why it is necessary to estimate user's intention.

Our proposal uses interactive method for user intention estimation. A local map is shown in system interface. A red arrow in the interface represents the orientation of the joystick on wheelchair. Once an object is detected as the destination, the object will be marked by green frame. At this moment, if the user holds the joystick pointing toward the object, the system will understand that the user intends to approach and the green frame will turn red to feedback to the user. This process is called target selection. Otherwise, pointing away from the object or releasing the joystick mean rejecting target and treating the object as obstacle, just as the previous shared control.

Motion Control. Target selection results in two effects. One is replacing the user commands by motion control commands to control the wheelchair automatically. The user's order through joystick is considered as user's intention on destination rather than direct control for velocity of the wheelchair. The velocity is calculated by motion control according to the position and orientation of the wheelchair and the goal. The other is the modification on parameters of shared control. For example, the threshold of obstacle avoidance is decreased to succeed passing through the narrow door or to perform fine manipulation; increase the obedience indicator to improve the accuracy of the tracking trajectory of the wheelchair.

We use a real-state feedback controller [13] to calculates the linear and angular velocity according to the relative position of the wheelchair and the target (Eq. 9).

$$\begin{cases} \mathrm{v} = k_r r \\ \omega = k_a a + k_b b \\ \mathrm{r} = \sqrt{\Delta x^2 + \Delta y^2} \\ a = -\theta + \mathrm{atan2}(\Delta x, \Delta y) \\ b = -\theta - a \end{cases} . \tag{9}$$

where Δx and Δy are position difference between the wheelchair and the target, θ is the orientation of wheelchair. The three variables above are with reference to the world coordinate system. k_r, k_a and k_b are constants.

5 Experiment and Results

Wheelchair Prototype. The wheelchair prototype [8] as shown in Fig. 2, is based on an ordinary electric wheelchair. The wheelchair is equipped with mobile robot sensors including LMS200 laser range finder (LRF), Kinect, odometry etc. A computer running Linux is used to implement the proposed shared control and 3D semantic mapping algorithm. A Smart Motion Controller (SMC) based on a DSP processor is adopted to execute motion control commands for wheelchair. Kinect is connected to the computer for obtaining 3D point cloud. The system software is developed based on ROS [14] and PCL [15].

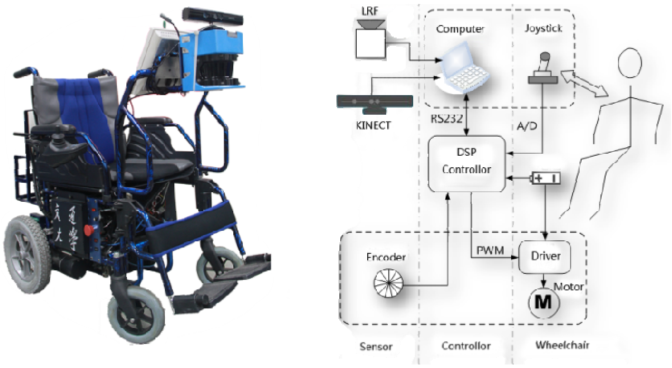


Fig. 2. Wheelchair prototype

Experimental Environment and Task. The experiment is implemented in the laboratory environment as shown in Fig. 5(a). The tasks of the experiments are driving the wheelchair starting from the passageway, then passing through a doorway to get into the laboratory, and finally docking into the table.

The blue line in Fig. 5(a) is the trajectory recorded by the odometry of wheelchair. Dashed part means that the wheelchair was controlled by motion control when the 3D semantic map detected the target and user intention selected the target. And the solid part means that the wheelchair is in the manual mode.

3D Semantic Map Building. Fig. 3 illustrates how 3D semantic map work with a table as example. The first column is the raw data obtain by Kinect. The next column is the 3D semantic map with parts of intermediate results, such as marking the surface of table and the wall, projecting the obstacle on the ground. The third column is the user interface shown on the PC which marks the target with green frame before select it. The last column shows the user interface after user select the target.

Table 1 shows the processing time for each step of map building in Fig. 3. It can satisfy the requirement of real-time of navigation system basically.

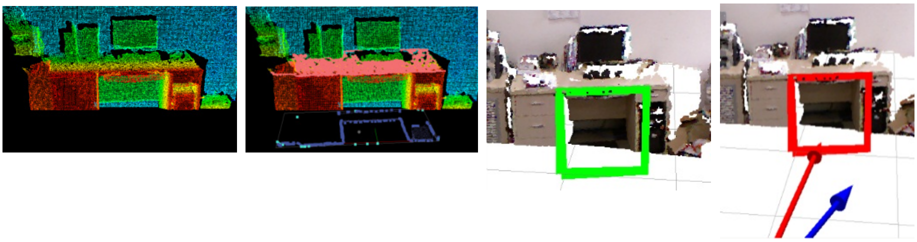


Fig. 3. D semantic map building and target selection

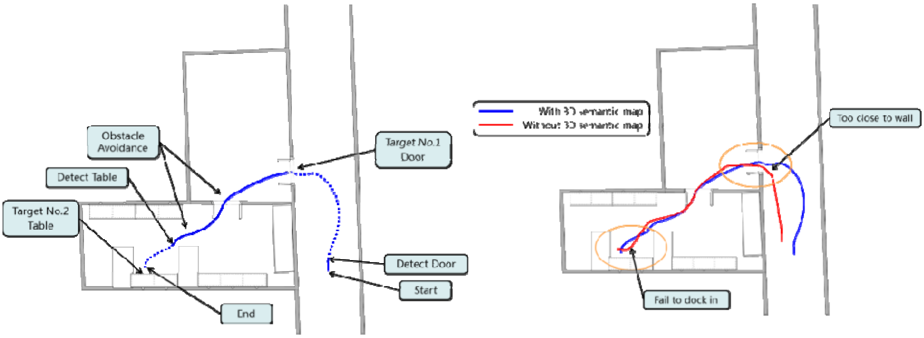
Table 1. Processing time for each step

	Table	Door
Down Sample	0.23s	0.23s
Segmentation	0.01s	0.05s
Model Matching	0.15s	0.96s
Detail Information Extraction	0.39s	0.02s
Total	0.78s	1.26s

Shared Control. Fig. 4 illustrates the accuracy of our control system. The width of wheelchair is 0.5m. It can be docked into the table with a 0.67m width free space.

Fig. 4. Docking into the table

Fig. 5(b) shows the comparison of trajectories with and without the 3D semantic map. As shown, when passing through the doorway, the wheelchair controlled with 3D semantic map took an arc to align the center of the door and passed through the doorway vertically. The wheelchair controlled without 3D semantic map, however, passed very close to one side, which is very dangerous. When docking into the table, the wheelchair controlled with 3D semantic map docked autonomously and precisely. On the opposite, the wheelchair couldn't approach to the table and failed to dock.



(a) (b)

Fig. 5. (a) Experimental environment and task; (b) Comparative experiment

Fig. 6 illustrates the three indicators and user weight for every moment in the trajectory with and without 3D semantic map. In order to show more clearly, the data in the figure is $(1 - value)$. The comparison declares that all indicators have been improved. The wheelchair will switch to automatic mode after target selection. The parameter of *safety* is reduced and the linear and angular velocity is calculated by machine, so the *safety* and *smooth* indicators are improved. Therefore, the *obedience* indicator also has been improve.

Fig. 7(a) shows the relationship between the user weight and three indicators and the objective function $\min(safety, comfort, obedience)$ when the wheelchair was in the location of the first point marked as "obstacle avoidance" in Fig. 5(a). Fig. 7(b) shows the Golden section search algorithm search process. Algorithm was converged after 18 steps, and error was less than 0.0001. It illustrates that the search interval gradually converge to the maximum value of the objective function fast and precisely.

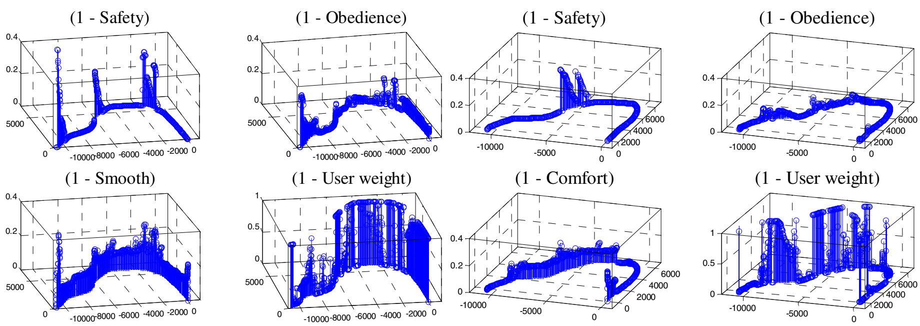
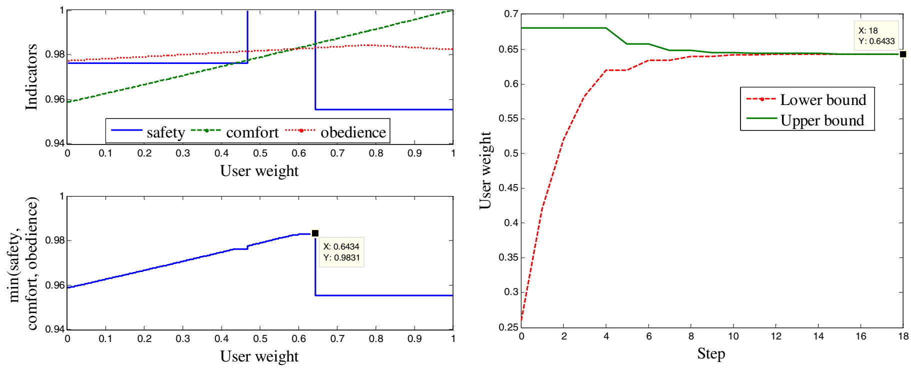


Fig. 6. Comparison of indicators of shared control with (right) and without (left) 3D semantic map



(a) (b)

Fig. 7. (a) Diagram of the indicators changes with user weight; (b) Convergence process of search algorithm

6 Conclusion and Future Works

This paper presents a 3D semantic map based-shared control for smart wheelchair. 3D semantic map is used to enhance the environment perception of wheelchair. The wheelchair is able to recognize different objects in unknown indoor environments, and with this information the wheelchair can assist the use to implement object related navigation tasks such as door passage or furniture docking. Further, the cooperation between human and wheelchair is improved based on the map. The experiments with real wheelchair and in real world illustrate the validity of the proposed method. In the future, the robustness and stability of the system for more complex environments will be further investigated.

Acknowledgments. This work is partly supported by the National High Technology Research and Development Program of China under grant 2012AA041403, the Natural Science Foundation of China under grant 60934006 and 61175088, the State Key Laboratory of Robotics and System (HIT), the Research Fund for the Doctoral Program of Higher Education under grant 20100073110018.

References

1. Bourhis, G., Agostini, Y.: Man-machine cooperation for the control of an intelligent powered wheelchair. Journal of Intelligent & Robotic Systems 22(3), 269–287 (1998)
2. Lankenau, A., Rofer, T.: A versatile and safe mobility assistant. IEEE Robotics & Automation Magazine 8(1), 29–37 (2001)
3. Parikh, S.P., Grassi Jr., V., Kumar, V., et al.: Integrating human inputs with autonomous behaviors on an intelligent wheelchair platform. IEEE Intelligent Systems 22(2), 33–41 (2007)

4. Carlson, T., Demiris, Y.: Human-wheelchair collaboration through prediction of intention and adaptive assistance. In: IEEE International Conference on Robotics and Automation, Pasadena, CA (2008)
5. Demeester, E., Hüntemann, A., Vanhooydonck, D., et al.: User-adapted plan recognition and user-adapted shared control: A Bayesian approach to semi-autonomous wheelchair driving. Autonomous Robots 24(2), 193–211 (2008)
6. Poncela, A., Urdiales, C., Pérez, E.J., et al.: A new efficiency-weighted strategy for continuous human/robot cooperation in navigation. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans 39(3), 486–500 (2009)
7. Urdiales, C., Peula, J., Barrue, C., et al.: An adaptive scheme for wheelchair navigation collaborative control. In: Association for the Advancement of Artificial Intelligence. Fall Symposium, Washington (2008)
8. Li, Q., Chen, W., Wang, J.: Dynamic shared control for human-wheelchair cooperation. In: IEEE International Conference on Robotics and Automation, Shanghai, China (2011)
9. Bailey, M., Chanler, A., Maxwell, B., et al.: Development of Vision-Based Navigation for a Robotic Wheelchair. In: IEEE 10th International Conference on Rehabilitation Robotics, Noordwijk, Netherlands (2007)
10. Murarka, A., Kuipers, B.: A stereo vision based mapping algorithm for detecting inclines, drop-offs, and obstacles for safe local navigation. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO (2009)
11. Rusu, R.B., Marton, Z.C., Blodow, N., et al.: Towards 3D Point cloud based object maps for household environments. Robotics and Autonomous Systems 56(11), 927–941 (2008)
12. Schnabel, R., Wahl, R., Klein, R.: Efficient RANSAC for Point-Cloud Shape Detection. Computer Graphics Forum 26(2), 214–226 (2007)
13. Siegwart, R., Nourbakhsh, I.R.: Introduction to autonomous mobile robots. The MIT Press (2004)
14. Robot Operating System (ROS), `http://www.ros.org/`
15. Rusu, R.B., Cousins, S.: 3D is here: Point Cloud Library (PCL). In: IEEE International Conference on Robotics and Automation, Shanghai, China (2011)