3D Visibility Check in Webots for Human Perspective Taking in Human-Robot Interaction

Ji-Hyeong Han and Jong-Hwan Kim

Department of Electrical Engineering, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, 305-701, Republic of Korea
{jhhan,johkim}@rit.kaist.ac.kr

Abstract. The rapid development of intelligent robotics would facilitate humans and robots will live and work together at a human workspace in the near future. It means research on effective human-robot interaction is essential for future robotics. The most common situation of human-robot interaction is that humans and robots work cooperatively, and robots should give proper assistance to humans for achieving a goal. In the workspace there are several objects including tools and a robot should identify the human intended objects or tools. There might be situational differences between a robot's perspective and a human perspective because of several obstacles in environment. Thus, a robot needs to take the human perspective and simulates the situation from the human perspective to identify the human intended object. For human perspective taking, first of all a robot needs to check its own visibility for the environment. To address this challenge, this paper develops a 3D visibility check method by using a depth image in Webots. By using the developed method, a robot can determine whether each point in the environment is visible or invisible at its posture and detect objects if they are visible.

Keywords: Human-robot interaction, 3D visibility check, human perspective taking, Webots.

1 Introduction

Since robot technology and intelligence technology have been matured, robots will come into our daily lives in the near future. Thus, the effective human-robot interaction (HRI) is needed especially from human-robot cooperation point of view. The research dealing with HRI problems has been intensively studied in various applications [1]-[7].

Among the various HRI problems, the human-robot cooperation problem should be solved first since getting a robot into a human workspace is the most possible and helpful way to get robots involved in human life. When a robot and a human work cooperatively, they should share the context information to achieve the goal successfully. The usual human-robot cooperation situation is that a robot assists a human partner and they work together using some tools or objects. Therefore, a robot needs to identify the human intended object to provide the proper assistance to a human.

J.-H. Kim et al. (eds.), *Robot Intelligence Technology and Applications 3*,
Advances in Intelligent Systems and Computing 345, DOI: 10.1007/978-3-319-16841-8_24

There might be situational differences between a robot's perspective and a human perspective because of several obstacles in environment; therefore, a robot should take the human perspective and consider the situation from the human perspective to identify the human intended object [8]. There can be four kinds of object states based on the robot's perspective and human perspective: i) an object that is visible from both perspectives, ii) an object that is visible from a robot's perspective and invisible from a human perspective, iii) an object that is invisible from a robot's perspective and visible from a human perspective, or iv) an object that is invisible from both perspectives. Among the objects of all the cases, the objects of the first and third cases might be the human intended objects since the object that a human can not see might not be the human intended one. Therefore, to identify the human intended object, first of all a robot needs to check its own visibility for the objects and also needs to check the human visibility by taking the human perspective.

To deal with the above mentioned issue, this paper develops a 3D visibility check method. The developed method considers 3D visibility rather than 2D visibility since a robot needs to consider the heights of obstacles. We develop the 3D visibility check method by using an RGB-D sensor in Webots. Webots is a robot simulator that has been widely used in robotics since 1998 [9]. In recent years, an RGB-D sensor has been used to wide robotics field for 3D modeling of environments, object and gesture recognition, simultaneous localization and mapping (SLAM), etc [10]-[12]. By using the proposed method, a robot can check which points in the environment are visible or invisible such that it can detect the objects if they are visible.

This paper is organized as follows. Section 2 presents the developed 3D visibility check method. In Section 3, experimental results are discussed and concluding remarks follow in Section 4.

2 The Developed 3D Visibility Check Method

In this section, the developed 3D visibility check method in Webots is explained along with step by step calculation. The developed method uses a depth image from an RGB-D camera sensor which locates at the top of a simulated mobile robot in Webots.

2.1 The Developed 3D Visibility Check Method

Fig. 1 shows coordinates of a depth image obtained from an RGB-D camera sensor in Webots and relationship between a depth image and real world. Since an RGB-D camera in Webots is a standard pinhole camera and it is implemented using OpenGL, the relationship between a depth image and real world can be modeled. The pixels in a depth image mean they are visible from the robot's perspective; thus, a robot can determine whether each point in real world is visible or invisible from its posture by using the relationship between the depth image and real world. The detailed calculations are explained in the following.

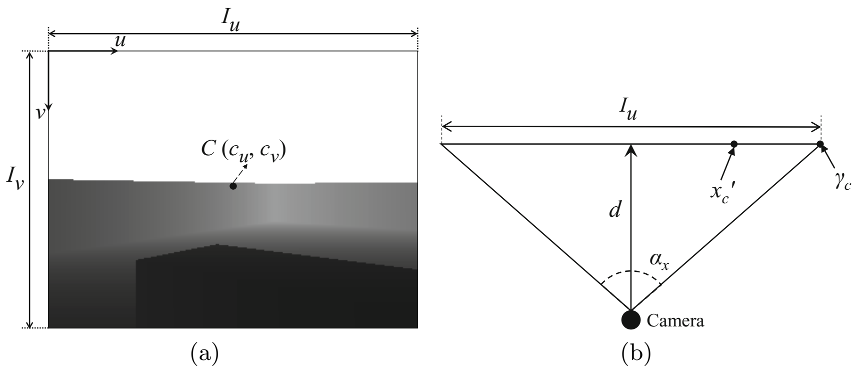


Fig. 1. (a) A depth image obtained from an RGB-D sensor in Webots. I_u and I_v are width and height of the image, u and v are pixel coordinates of the image, and $C(c_u, c_v)$ is a center of the image. (b) The relationship between a depth image in u-coordinate and real world in camera's x-coordinate. α_x is a horizontal field of view angle of the camera, d is a depth value obtained from a depth image, and x'_c and γ_c are points of the real world in camera's x-coordinate.

The relationship between a pixel of a depth image, $p(u', v')$ and a corresponding point of real world based on a camera posture, $q_c(x'_c, y'_c, z'_c)$ can be represented as a proportional expression. The relationship between u-coordinated of a depth image and x-coordinate in real world camera posture is represented as follows:

$$\gamma_c : \frac{I_u}{2} = x'_c : (u' - c_u) \tag{1}$$

where u' is a pixel of a depth image in u-coordinate, x'_c is a corresponding point of real world in camera's x-coordinate, I_u is a width of a depth image, and $c_u = I_w/2$. γ_c is a point of real world in camera's x-coordinate corresponding to a right-hand end pixel of a depth image in u-coordinate and it is represented as follows:

$$\gamma_c = d \tan \frac{\alpha_x}{2} \tag{2}$$

where d is a depth value of p obtained from a depth image and α_x is a horizontal field of view angle of the camera. Therefore, a point of real world in camera's x-coordinate is represented as follows:

$$x'_c = \gamma(u' - c_u)\frac{2}{I_u} = \frac{2d(u' - c_u)\tan(\alpha_x/2)}{I_u}. \tag{3}$$

In the same manner as the calculation of x'_c, y'_c, which is a point of real world in camera's y-coordinate corresponding to a bottom end pixel of a depth image in v-coordinate, is represented as follows:

$$y'_c = \frac{2d(v' - c_v)\tan(\alpha_y/2)}{I_v} \tag{4}$$

where v' is a pixel of a depth image in v-coordinate, I_v is a height of a depth image, $c_v = I_v/2$, and $\alpha_y = \alpha_x \frac{I_v}{I_u}$ is a vertical field of view angle of the camera. A point of real world in camera's z-coordinate is

$$z'_c = d. \tag{5}$$

By using the above calculation, a robot can match a pixel of a depth image in (u, v)-coordinate to a point of real world in camera's (x, y, z)-coordinate.

The next step is transforming $q_c(x'_c, y'_c, z'_c)$ to $q_g(x'_g, y'_g, z'_g)$ that is based on global (x, y, z)-coordinate. To transform the camera's coordinate to the global coordinate, the transformation matrices are calculated using the relative posture of the camera to the robot and the relative posture of the robot to the world in Webots. The transformation is done as follows:

$$\mathbf{q}_g = {}^{g}_{r}\mathbf{T}\,{}^{r}_{c}\mathbf{T}\,\mathbf{q}_c \tag{6}$$

where ${}^{r}_{c}\mathbf{T}$ is a transformation matrix from camera's coordinate to robot's coordinate and ${}^{g}_{r}\mathbf{T}$ is a transformation matrix from robot's coordinate to global coordinate.

Objects in the environment can be detected if a robot identify they are visible at robot's posture by using the developed 3D visibility check method. Each point in the environment is under three cases. The first case is that the point has a depth value exceeding the maximum depth range of an RGB-D camera. In this case, the point is considered as empty like the sky and ignored. The second case is that y-value of the point is close to zero, i.e. $y < \sigma$, where σ is a small value like 1.0 cm. It means that the point belongs to the ground and the reason using σ is that an RGB-D camera has noise. The last case is that the point is a part of an object, when satisfying $||\mathbf{C}_{obj} - \mathbf{P}||_2 < ObjectSize/2$, where $\mathbf{C}_{obj}$ is a center point of an object and $\mathbf{P}$ is a considered point.

3 Experiment

This section presents the experimental environment and results that were conducted to show the effectiveness of the developed method.

3.1 Experimental Environment

A differential wheel robot with an RGB-D camera sensor at its top was used to show the effectiveness of the developed method. Fig. 2 shows the experimental environment. There were two obstacles with different heights, one object (a ball), and a simulated human. The height and width of ground were 4.0 m. The parameters of the RGB-D camera mounted on the robot were a horizontal field of view angle as 0.994837 radian, width as 320, height as 240, minimum and maximum depth range as 0.3 m and 4.5 m, respectively, in the same manner as Kinect.

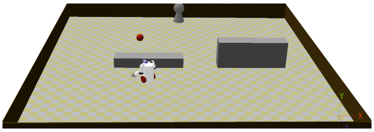
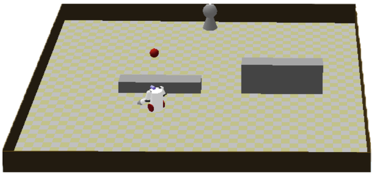
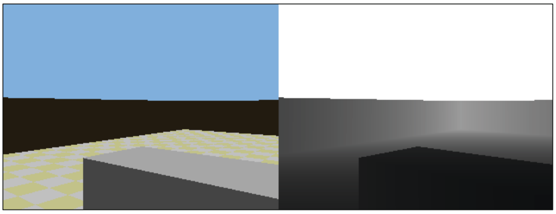


Fig. 2. The experimental environment in Webots



(a) An experiment scene



(b) RGB and depth images obtained from an RGB-D camera.

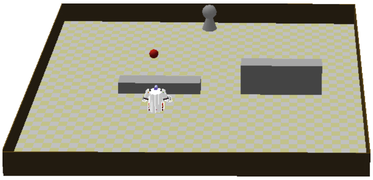
(c) 3D visibility check result

Fig. 3. Results of experiment 1

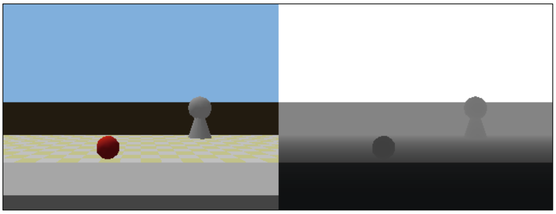
3.2 Experimental Result

To show the effectiveness of the developed method, four different experiments with different robot's postures and object positions were conducted. Figs. 3-6 show experiment scenes, RGB and depth images obtained from an RGB-D camera mounted on the robot, and 3D visibility check results of experiments 1, 2, 3, and 4, respectively. The 3D visibility check result figures are projections of 3D points of real world to 2D plane to show the visibility check result effectively. In the figures of 3D visibility check result, the circle with gray is a robot, black means invisible area, green means visible area, brown means the visible upper part of obstacles, red means a visible object, and blue means a visible human.

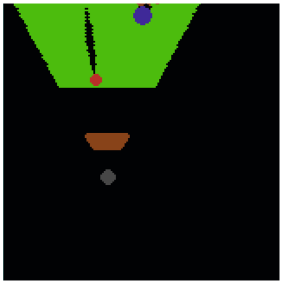
As shown in the figures, the robot could check its 3D visibility well and detect the object and human if they are visible. Also, the robot could identify the empty space like the sky and ground as well. The robot recognized that the area behind the obstacles or objects is invisible. By the developed 3D visibility check rather than 2D visibility check, the robot could recognize the situation difference caused



(a) An experiment scene

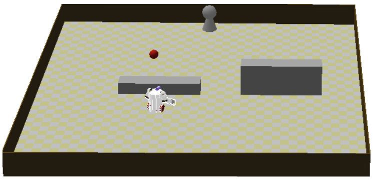


(b) RGB and depth images obtained from an RGB-D camera.

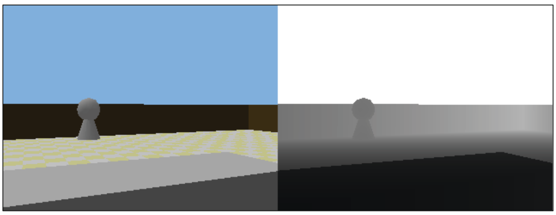


(c) 3D visibility check result

Fig. 4. Results of experiment 2



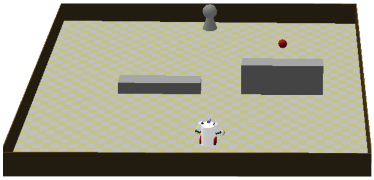
(a) An experiment scene



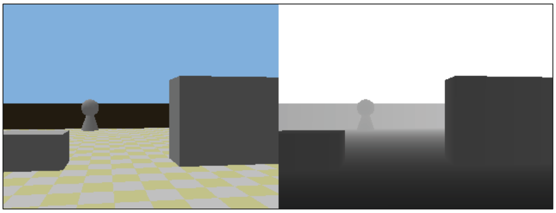
(b) RGB and depth images obtained from an RGB-D camera.

(c) 3D visibility check result

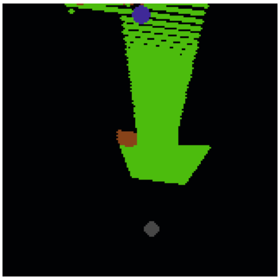
Fig. 5. Results of experiment 3



(a) An experiment scene



(b) RGB and depth images obtained from an RGB-D camera.



(c) 3D visibility check result

Fig. 6. Results of experiment 4

by heights of obstacles as shown in Fig. 6(c). In Fig. 6(c), a robot identified that the upper part of lower obstacle is visible, but the upper part of higher obstacle is invisible.

4 Conclusion

This paper presents the development of a 3D visibility check method using a depth image in Webots. By using the relationship between a depth image and real world, a robot could identify whether each point in real world is visible or invisible at its posture. Further more, a robot could detect objects based on the result of visibility check. The developed method dealt with 3D visibility rather than 2D visibility to consider the heights of obstacles and objects. This method can be used to make a robot take a human perspective for effective human-robot interaction. Because a robot needs to determine its own visibility first and then take a human perspective by simulating the situation assuming that a robot is located at human posture and calculating the human visibility in the same manner as its own visibility calculation. As further research, we will make a robot find out the human intended object by human perspective taking even when there are situational difference between robot's perspective and human perspective.

Acknowledgment. This work was supported by the Technology Innovation Program, 10045252, Development of robot task intelligence technology, funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea).

References

1. Arkin, R.C., et al.: An ethological and emotional basis for human-robot interaction. Robotics and Autonomous Systems 42(3), 191–201 (2003)
2. Ros, R., et al.: Adaptive human-robot interaction in sensorimotor task instruction: From human to robot dance tutors. Robotics and Autonomous Systems 62(6), 707–720 (2014)
3. Severinson-Eklundh, K., et al.: Social and collaborative aspects of interaction with a service robot. Robotics and Autonomous Systems 42(3), 223–234 (2003)
4. Sorbello, R., et al.: Telenoid android robot as an embodied perceptual social regulation medium engaging natural human-humanoid interaction. Robotics and Autonomous Systems (2014), http://dx.doi.org/10.1016/j.robot.2014.03.017
5. Schmidt, P.A., et al.: A sensor for dynamic tactile information with applications in human-robot interaction and object exploration. Robotics and Autonomous Systems 54(12), 1005–1014 (2006)
6. Fritsch, J., et al.: Multi-modal anchoring for human-robot interaction. Robotics and Autonomous Systems 43(2), 133–147 (2003)
7. Cifuentes, C.A., et al.: Human-robot interaction based on wearable IMU sensor and laser range finder. Robotics and Autonomous Systems (2014), http://dx.doi.org/10.1016/j.robot.2014.06.001
8. Trafton, J.G., et al.: Enabling effective human-robot interaction using perspective-taking in robots. IEEE Trans. Systems, Man, and Cybernetics-Part A: Systems and Humans 35(4), 460–470 (2005)
9. Webots: robot simulator, http://www.cyberbotics.com
10. Henry, P., Krainin, M., Herbst, E., Ren, X., Fox, D.: RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments. In: Khatib, O., Kumar, V., Sukhatme, G. (eds.) Experimental Robotics. STAR, vol. 79, pp. 477–491. Springer, Heidelberg (2012)
11. Ramey, A., et al.: Integration of a low-cost RGB-D sensor in a social robot for gesture recognition. In: Proc. 6th International Conference on HRI, pp. 229–230 (2011)
12. Hu, G., et al.: A robust RGB-D slam algorithm. In: Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1714–1719 (2012)