Object Detection for a Mobile Robot Using Mixed Reality

Hua Chen[1], Oliver Wulf[2], and Bernardo Wagner[2]

[1] Learning Lab Lower Saxony (L3S) Research Center, Hanover, Germany
chen@l3s.de

[2] Institute for Systems Engineering, University of Hanover, Germany
{wulf,wagner}@rts.uni-hannover.de

Abstract. This paper describes a novel Human-Robot Interface (HRI) that uses a Mixed Reality (MR) space to enhance and visualize object detection for mobile robot navigation. The MR space combines the 3D virtual model of a mobile robot and its navigating environment with the real data such as physical building measurement, the real-time acquired robot's position and laser scanned points. The huge amount of laser scanned points are rapidly segmented as belonging either to the background (i.e. fixed building) or newly appeared objects by comparing them with the 3D virtual model. This segmentation result can not only accelerate the object detection process but also facilitate the further process of object recognition with significant reduction of redundant sensor data. Such a MR space can also help human operators realizing effective surveillance through real-time visualization of the object detection results. It can be applied in a variety of mobile robot applications in a known environment. Experimental results verify the validity and feasibility of the proposed approach.

1 Introduction

Last two decades have seen the popularization of Virtual Reality (VR) that attempts to use virtual environments to convey a sense of reality. A virtual environment is the representation of the computational geometric modeling of a real environment in a 3D visual simulation created with computer graphics techniques [1]. Since real objects that are difficult to access can be represented by virtual objects in a virtual environment, VR finds applications in many fields such as military training, medicine and education etc. Especially, VR has huge potential in robotics [2] and it has facilitated the development of new kind of Human-Robot Interface (HRI). Through VR, human operators are able to interact with the real robots that work in a remote, hazardous, micro or macro environment in an intuitive, natural way.

According to the correlation between the real world and the virtual environment, the VR-based HRI tools fall into two main categories. One is using virtual environments as a front-end interface for controlling a robot such as the applications in telerobotic manipulation [3,4]. Physical robots and their working space are represented in the virtual environments that provide visual cues,

H. Zha et al. (Eds.): VSMM 2006, LNCS 4270, pp. 466–475, 2006.

constraints or predicted trajectory for more accurate and safe execution of the operator's commands. The other kind of HRI constructs virtual environments to simulate the behavior and working space of a robot. In this case, the physical robot and its surroundings are not even required. Currently, robotic system is still expensive and fragile, which may preclude most students from having "hand-on" experience and also prevents researchers from trying any robotic algorithm freely because of some safety issues. Therefore, virtual environments used for simulations can make robotics not only more accessible to students as a learning tool [5] but also more flexible to researchers as a testbed [6]. Furthermore, it is necessary to use virtual environments to simulate some specific environments such as disaster areas for the investigation on robots that undertake the Urban Search-And-Rescue (USAR) tasks [7].

Recently, the importance of developing a HRI that meets the users' needs and requirements has received more attention [8]. In semi-autonomous mobile robots applications, some real information such as paths to survivor [10] or navigation map [11] is visualized in a virtual environment using Mixed Reality (MR) technique. Introduced by [17], MR technique aims at making use of both the flexible computer modeling and rich knowledge information in the real world within a MR space. It has received much interest due to its potential of helping humans in some practical activities such as industrial design [9]. This kind of MR-based HRI assists human operators in situation awareness, monitoring and control for surveillance during the collaboration with mobile robots. This paper addresses a novel MR-based HRI for mobile robot applications.

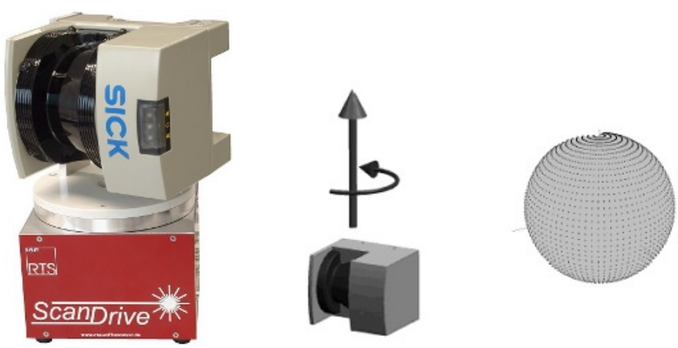
Given a prior known environment, indoors as well as outdoors, a MR space is constructed to enhance and visualize the object detection during mobile robot navigation. Here the object is defined as the newly appeared object in the environment. The MR space integrates the 3D virtual model of a robot and its known navigating environment with the real data such as the physical building measurement, the real-time acquired robot's position and laser range data. By comparing with the 3D virtual model, the huge amount of laser range data is rapidly segmented as belonging either to the known background (i.e. fixed building) or newly appeared objects. This segmentation result can not only speed up the object detection process but also facilitate the following object recognition process with significant reduction of redundancy in sensor data. The MR space provided in the proposed HRI can improve the robot intelligence on object detection through the comparison with the virtual model and it can also help human operators realizing effective surveillance through real-time visualization of the detection results. Such a HRI tool can be applied in a variety of mobile robot applications such as exploring indoor environments (e.g. office/house buildings) to detect victims, survivors and fallen obstacles for rescue task or navigating outdoor environments (e.g. a parking square) to detect occupied spaces and further find spaces for parking.

This paper is organized as follows. Section 2 introduces the sensor system that is used to acquire range data in our experiments. Section 3 describes the method of object detection in detail. Experimental results are given in section 4. Finally, the conclusions and some future work are summarized in section 5.

2 Range Data Acquisition

Range data that indicate the distances from obstacles to a robot are inherently suited for the object detection task. Based on different sensor systems, various methods have been provided to acquire range data. Sonar sensors use the transmission and reflection of sound wave to find obstacles [12], but they suffer from poor directionality, specular reflections and wrong distance measurement caused by crosstalk. A camera-stereo pair can be used to derive 3D information from a calculated disparity map [13], but the captured image is sensitive to light conditions (e.g. shadows).

Currently, laser scanners that provide better resolution are getting used widely within the robotics community. Among various laser scanners, the laser scanner based on the measurement of Time-Of-Flight (TOF) is most suitable for fast scanning during mobile robot navigation and it is used in this work. The distance between the scanner and the object surface is measured by the time between the transmission and the reception of the laser beam. Many commercial 2D TOF laser scanners are available for the distance measurement on one plane.



(a) (b)

Fig. 1. The solution to 3D scanning (a)A 3D scanning system consisting of a 2D scanner and a servo drive (b)The yawing scan pattern

For 3D object detection, a laser scanner that can measure distance in 3D space is desired. However, there are no commercial 3D scanners available for this purpose at present. In our system, a commercial 2D scanner and a mechanical actuator are combined to achieve 3D scanning (see Fig. 1a). The actuator is a self-build servo drive that turns the 2D scanner around the upright axis (i.e. yawing scan pattern) to reach the third dimension (see Fig. 1b).

The usability and reliability of a 3D scanning system are evaluated by some main criteria such as the ability of fast and continuous scanning while driving, accurate synchronization of the laser measurement and the scanning device etc. After our previous optimization work [15] such as mechanical improvements of the scanning device and compensation of systematic measurement errors, this 3D scanning system can be used to acquire an undistorted 3D point cloud in a range about 30m with centimeter accuracy. In our experiments, during the navigation of a robot equipped with this scanning system, the undistorted 3D point cloud is acquired at intervals of 2.4s with 2^o and 1^o resolution at horizontal and vertical directions, respectively.

3 Object Detection in a MR Space

With prior knowledge of a mobile robot navigating environment, objects here are defined as newly appeared objects in the known environment. In this section, the construction of a 3D MR space corresponding to a robot navigating environment is firstly described. Further, the object detection within the MR space is presented.

3.1 MR Space Construction

Here the MR space is constructed by combing the virtual model of a mobile robot and its navigating environment with the real data such as the physical building measurement, the real-time acquired robot's position and laser scanned points.

Firstly, a static 3D virtual scene model corresponding to the physical mobile robot navigation is built using some popular 3D design tool (e.g. 3D MAX). The size of the 3D model is appropriately scaled to ensure the desired quality and speed of 3D rendering. Then the designed 3D model is exported to a Virtual Reality Modeling Language (VRML) file that is a ratified International Organization for Standardization (ISO) standard (ISO 14772) file format for description of the geometry and behavior of 3D computer graphics. In a VRML file, a virtual scene is described by a scene graph in which virtual objects are represented as VRML nodes and organized in hierarchical structure. Since a web browser with a VRML plug-in is able to show a VRML-based virtual scene by interpreting the corresponding VRML file, VRML is largely used to distribute 3D applications via the World Wide Web and it is widely supported by current prevalent 3D design software. Here the static VRML-based virtual scene is composed of a virtual robot and a virtual building model that correspond to the physical robot and its navigating environment respectively.

During the mobile robot navigation, the robot pose (i.e. position and orientation) is obtained using our novel robust localization method that integrates a 3D laser scanning into a Monte Carlo Localization (MCL) system [16]. Range data are acquired from the scanning systems described above. These dynamic real data are required to be integrated with the virtual scene to form the dynamic MR space. This integration is accomplished by using Java Applet to communicate between the robot system and the MR space via VRML External Authoring

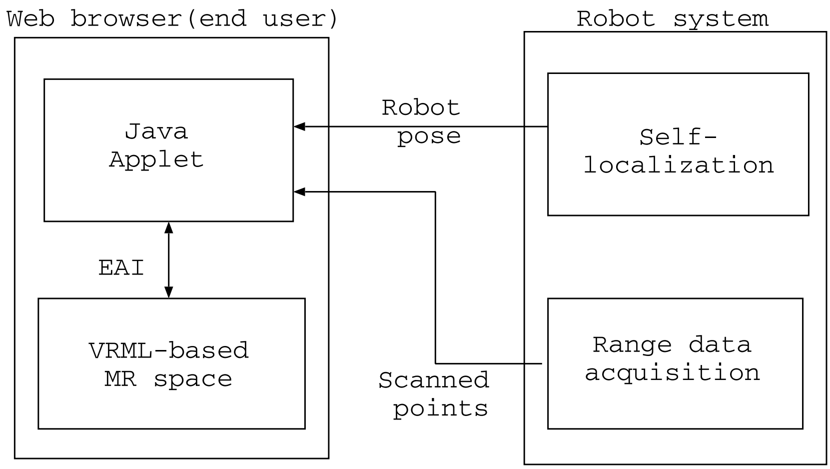


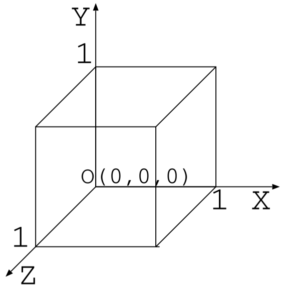
Fig. 2. The dynamic MR space interface architecture

Interface (EAI) that defines a set of functions to allow an external program to access VRML nodes (see Fig. 2).

A Java applet running in the end user's web browser needs to be firstly signed as a trusted applet for the permission of a safe connection with the robot system. Then the signed Java Applet receives the dynamic data produced by the navigating mobile robot in real-time via network using the TCP/IP protocol. These dynamic 3D data are transformed to be represented in an appointed world coordinate frame for data consistency. Finally, VRML EAI allows this external Java Applet to update the VRML-based MR space such as set the virtual robot to have the same pose with the real robot, render the MR space according to the current viewpoint of the real robot and create 3D graphic objects that represent laser scanned points. Note that creating and rendering the huge amount of 3D graphic objects is an extremely time-consuming process under current non-specialized hardware condition. One way to accelerate the rendering process is to preload the 3D graphic objects off-line and transform them to the corresponding positions in real-time during on-line process.

3.2 Object Detection

The object here is defined as the newly appeared object at some different position from a known robot's navigating environment. As described above, the robot's navigating environment can be represented by a virtual building model. Thus any scanned point that differs in position from the virtual building model is segmented as belonging to an object. To implement this segmentation process, a table can be built to pre-store the 3D measurement of a virtual building model. Each line in the table indicates a geometrical surface in the model. Take a unit cube as an example. The 3D data stored in a table represent six faces of the cube at each line respectively (see Fig. 3).



x	y	z
0	$[0, 1]$	$[0, 1]$
1	$[0, 1]$	$[0, 1]$
$[0, 1]$	0	$[0, 1]$
$[0, 1]$	1	$[0, 1]$
$[0, 1]$	$[0, 1]$	0
$[0, 1]$	$[0, 1]$	1

Fig. 3. (left) A unit cube model; (right) A table representing six faces of a unit cube

With the proposed table that represents all the geometrical surfaces of a model, the following steps can be used to determine whether a given point belong to an object or the model:

1. Compute the distances from the given point to the surfaces represented by each line in the table respectively
2. Find the minimum of the distances resulted from step (1)
3. The given point belongs to an object if the minimum value is above an appointed threshold

In ideal case, the given point belongs to an object if the minimal distance does not equal 0, which means it is not on any surface of the model. In practice, the noise, the measurement error and the accuracy of sensor data all need to be taken into account. So in step (3), the given point is determined to be on an object if it is enough far (more than a threshold value) away from the nearest surface of the model.

Based on the principle above, the object detection for a mobile robot can be performed by comparing the laser scanned points with the table that stores the 3D data of the virtual building model corresponding to the robot's navigating space. Such a table is built along with the design of the virtual environment. In practice, the real navigating environment may consist of a large number of complex geometrical surfaces, which will drastically increase the size of the table for 3D data storing. Comparing each scanned point with each surface represented by each line in the huge-sized table will slow down the detection process. Our solution is using a valid segment of the table for comparison. The valid segment indicates a certain part of the virtual building model corresponding to the nearest region that the sensor data can reach. It can be implemented based on the structured model and the recorded positions of scanned points related to the robot. For example, in our experiment, a forward scanned point only needs to be compared with the current forward-looking surfaces of the robot within its reachable range of 30m. Finally, the detection result is represented by the laser scanned points that belong to newly appeared objects within the MR space.

4 Experimental Results

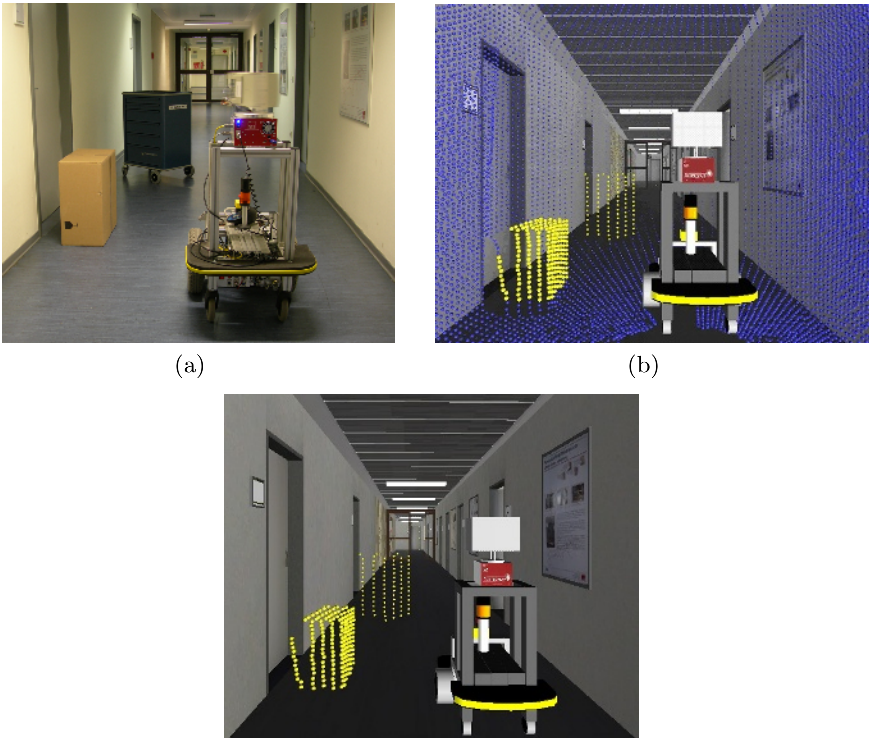
For experimental verification, the corridor of our institute has been selected as a typical indoor environment for robot navigation. A 3D virtual building model corresponding to the institute environment has been constructed. Two mobile robots that are equipped with a commercial 2D TOF laser scanner and the proposed 3D scanning system respectively were used to navigate along the corridor at the speed about 0.7m/s. The range data are continuously acquired during the robot navigation. Two virtual robots corresponding to these two mobile robots have also been built. At the user end, a HRI runs on a notebook PC (with a 1.69 GHz CPU and a 512 MB RAM) that is connected with the robot system via 11Mbit/sec Wireless Local-Area Network (WLAN). The HRI shows a MR space in a web browser (e.g. Internet Explorer) with a VRML plug-in (e.g. Cortona [14]). Fig. 4 and Fig. 5 give the experimental results of a robot with the 2D scanner and the 3D scanning system respectively.

(a) (b)

Fig. 4. Robot navigation with a 2D laser scanner (a) A HRI corresponding to the physical navigation environment (b) The HRI shows a MR space where detected objects are highlighted in yellow

Fig. 4(a) presents the physical corridor environment where a robot with a 2D scanner is navigating and the corresponding MR space displayed on the notebook PC as a HRI for the user. Fig. 4(b) demonstrates a virtual robot navigating in the MR space, which corresponds to the physical robot navigation. The position and orientation of the virtual robot are the same with those of the real robot. The 2D scanned points are visualized with 3D graphic objects (i.e. small cylinders) at corresponding positions in the MR space. Here about 720 scanned points are acquired for a 360^o scan on one horizontal plane with 0.5^o resolution. Scanned points belonging to objects, i.e. two newly appeared boxes in the corridor, are highlighted with yellow color to show the detection result.

Fig. 5(a) gives the physical navigating environment of a robot with a 3D scanning system. Fig. 5(b) visualizes the 3D scanned point cloud with 3D graphic objects (i.e. small spheres) in the MR space. The highlighted scanned points in



(a) (b)

(c)

Fig. 5. Robot navigation with a 3D laser scanning system (a) Physical navigation environment (b) The 3D scanned point cloud is visualized in the MR space where the detected objects are highlighted in yellow (c) Only the scanned points belonging to the detected objects are visualized in the MR space

the MR space indicate the detected newly appeared objects. Fig. 5(c) visualizes the scanned points that only belong to the detected objects in the MR space.

Note that the amount of the 3D scanned points acquired by aforementioned 3D scanning system with 2^o and 1^o resolution at horizontal and vertical direction respectively is actually huge (about 32400). It is still difficult to render the MR space with so huge amount of 3D graphic objects in real-time with current non-specialized hardware. However, the MR space rendering can be substantially accelerated without visualizing the large amount of redundant scanned points as shown in Fig. 5(c). In our experiments, the system is able to visualize the detection result within the MR space in real-time. That is, during the mobile robot navigation, the virtual corridor, the virtual robot and 3D graphic symbols of laser scanned points belonging to newly appeared objects are dynamically updated in the MR space at each interval of laser scanning.

5 Conclusions

This paper presents a novel HRI that utilizes a MR space to enhance and visualize the object detection for mobile robot navigation in a known environment. This kind of HRI has dual functions. One is that the 3D data of the virtual environment corresponding to a real known navigating space of a mobile robot can be used by the robot to improve its intelligence on fast detection of newly appeared objects, and the other is that the real-time visualization of the object detection result is beneficial to human operators for effective surveillance purpose. The huge amount of laser range data has been considerably reduced to the data that only belong to the newly appeared objects. Currently, the detected objects are represented by the laser scanned points. With these reduced laser scanned points, a fast reconstruction of the 3D surface corresponding to each object could be further investigated to facilitate real-time 3D object recognition during mobile robot navigation.

References

1. Y. Ohta and H. Tamura (eds.), Mixed Reality: Merging Real and Virtual Worlds, Springer Verlag, 1999.
2. E. Natonek, L. Flckiger, Th. Zimmerman and C. Baur, Virtual Reality: an intuitive approach to robotics, in Proc. of SPIE Telemanipulator and Telepresence Technologies, Vol. 2351, pp. 260-270, 1994.
3. L. B. Rosenberg, Virtual fixtures: Perceptual tools for telerobotic manipulation, in Proc. IEEE Virtual Reality Annual International Symposium (VRAIS), pp. 76-82, 1993.
4. A. Monferrer and D. Bonyuet, Cooperative robot teleoperation through virtual reality interfaces, in Proc. of the Sixth International Conference on Information Visualisation, pp. 243 - 248, 2002.
5. T. L. Dunn and A. Wardhani, A 3D Robot Simulation for Education, in ACM Proc. of the 1st International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia, pp. 277-278, 2003.
6. D. Gracanin, M. Matijasevic, N.C. Tsourveloudis, K.P. Valavanis, Virtual reality testbed for mobile robots, in Proc. of the IEEE International Symposium on Industrial Electronics (ISIE), Vol. 1, pp. 293 - 297, 1999.
7. J.J Wang, M. Lewis and J. Gennari, A game engine based simulation of the NIST urban search and rescue arenas, in Proc. of the Winter Simulation Conference, pp.1039-1045, 2003.
8. J. A. Adams, Critical Considerations for Human-Robot Interface Development, in Proc. of the American Association for Artificial Intelligence (AAAI) Fall Symposium on Human-Robot Interaction, pp.1-8, 2002.
9. Hua Chen, Real-time video object extraction for Mixed Reality applications, Dissertation, Technical University of Clausthal, ISBN 3-89720-844-x, Germany, 2005
10. M. Daily, Y. Cho, K. Martin and D. Payton, World embedded interfaces for human-robot interaction, in Proc. of the 36th Annual Hawaii International Conference on System Sciences, pp. 125-130, 2003.
11. S. Tejada, A.Cristina, P. Goodwyne and et al., Virtual Synergy: A Human-Robot Interface for Urban Search and Rescue, In Proc. of the American Association for Artificial Intelligence (AAAI) Mobile Robot Competition, pp. 13-19, 2003.

12. R. Kuc, A spatial sampling criterion for sonar obstacle detection, in IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, Issue 7, pp. 686-690, 1990.
13. K. Berns, V. Kepplin and R. Dillmann, Terrain and obstacle detection for walking machines using a stereo-camera-head, in IEEE Proc. of the 24th Annual Conference of the Industrial Electronics Society (IECON), Vol. 2, pp.1170-1175, 1998.
14. http://www.parallelgraphics.com/products/cortona/
15. O. Wulf and B. Wagner, Fast 3D-Scanning Methods for Laser Measurement Systems, in International Conference on Control Systems and Computer Science (CSCS), 2003.
16. O. Wulf, M. K. Allah and B. Wagner, Using 3D Data for Monte Carlo Localization in Complex Indoor Environments, 2nd European Conference on Mobile Robots (ECMR), 2005.
17. P. Milgram and F. Kishino, A Taxonomy of Mixed Reality Visual Displays, in IEICE Transactions on Information Systems, Vol. E77-D, No.12, pp.1321-1329, 1994.