Immersive virtual environment for mobile platform remote operation and exploration

Jonathan Fournier, Marielle Mokhtari, and Benoit Ricard
Defence R&D Canada - Valcartier
Quebec, QC, Canada
{jonathan.fournier, marielle.mokhtari, benoit.ricard}@drdc-rddc.gc.ca

Abstract—**These days, robotic platforms are commonly used in operational conditions where manned operations are not practical, not cost-effective or too dangerous. Those robotic devices rely heavily on remote operations using imagery acquired by on-board sensors that provide quite limited situational awareness to the user. In difficult scenarios, this lack of good situational awareness could lead to the failure of the mission. This paper presents a new concept currently in development that will improve situational awareness of the remote platform operator through an immersive virtual environment. The system uses an immersive chamber (CAVE) in which the operator is able to visualize and interact with an avatar of a robot evolving in a 3D model of its area of operation. The 3D model is incrementally built from the remote platform sensor feeds and provides "persistent data" to the user. This paper presents the first phase of the work which involves the development of a concept demonstration prototype. The implementation uses a robot simulator instead of a real world robot in order to rapidly be able to evaluate the concept and perform experiments. The tools developed in simulation will serve as the base for further developments and support the transition to a real robotic platform.**

Immersive environment; remote operation; mobile platform; CAVE; robotic; ISR; surveillance; reconnaissance

I. INTRODUCTION

For years now, mobile sensor platforms such as robots, unmanned ground vehicles or unmanned air vehicles, can be controlled by transmitting their navigation sensor signal to a remote operator. Video feed and telemetry are displayed at a platform control station where operators make decisions and direct the platform and its sensor package. When the platform operates in a low to moderate complexity environment this scheme of operation can be sufficient to gather required imagery and data. However, using direct sensor feeds to operate air or ground platforms in higher complexity environments may become rapidly challenging. In this type of environment, proximity of obstacles, communication lag, short interruptions in data/video transmission and the need to perform 3D navigation reduce the reaction time (from detection to action) and navigation complexity goes beyond the limits of mission security. A perception of increased complexity is introduced due to the poor situational awareness provided by the platform sensor feeds. Cameras provide perception capabilities limited to the field-of-view and aiming direction of the active camera resulting in large portions of the operational envelope of the platform that cannot be monitored during the mission's critical phases.

The main limiting factor when operating a robotic platform is the lack of good situational awareness due to the absence of *data persistence*. Persistence of data refers to the ability to use platform localization and attitude sensors to record and fix in time and space the acquired data from perception sensors. This concept means much more than recording and replaying imagery and its associated metadata. It means that acquired perception sensor data is stored, georeferenced and can be consulted back in time and from different perspectives. To be used as such, the *persistent data* has to be used in conjunction with a 3D data space where incoming data is either used to build a 3D model or located and projected in the data space. This allows the operator to navigate through the accumulated perception data in time and space to gain the situational awareness required for efficient perception, exploration and navigation of robotic platforms in highly complex environment. It is envisioned that the generation of persistent data would be particularly useful for missions such as bomb clearing operation or advanced search in rubbles, where the work environment is quasi-static.

This paper presents a concept demonstration project investigating a Human-centric immersive working environment, also called virtual work cell, for remote robot operation and exploration. The virtual work cell allows iterative acquisitions from the robot's sensors of a scalable and interactive immersive 3D model of the real environment. The virtual model can be refined locally and / or globally according to what the operator wants to focus on. The work cell captures the motions and commands of the immersed operator to control the robot's behaviour. The operator controls a platform evolving in a real environment through an immersive environment (Fig. 1).

Figure 1. General concept of the demonstration project

II. RELATED WORK

A. Control

Teleoperated mobile platforms are generally controlled using a joystick or a regular mouse [2, 3, 4]. Kadavasal and Oliver [1] introduce a hybrid control strategy where temporary autonomous operations of the platform are allowed for local obstacle avoidance. In addition to manual control, Simsarian [6] adds the possibility to select an object in the virtual environment as the next target position of the platform and provides a speech recognition interface to allow controlling the robot using natural language commands.

B. Sensing and Display

Most mobile platforms use video camera for environment perception. While live camera feeds can be displayed on a computer screen, in [7, 8], the authors have investigated the advantages of using stereoscopic visualization for mobile robot teleoperation. It is also possible to equip a platform with an omnidirectional camera and display the live video feed to the operator via a head mounted display (HMD). The HMD can be combined with an inertial sensor so as to enable the adjustment of the image according to the orientation of the operator's head [2, 10]. The platform can also be equipped with a sonar. Lin and Kuo [4] use this type of sensor for positioning a remote controlled underwater vehicle relative to a CAD model of the environment. Yu *and al.* [5] use the sonar readings to generate a 2D map of the surroundings of their robot. Sensors can also be combined to create a richer representation of the environment. Ferland *and al.* [9], use a stereo camera combined with a laser range finder to generate textured 3D models of the robot's environment. The operator can choose to visualize the 3D model either from a first person or a third person viewpoint.

III. CONCEPT DESCRIPTION

To increase the situational awareness when operating a robotic platform in an highly complex but quasi-static environment, the authors have designed a system that makes use of a virtual reality (VR) immersion chamber (CAVE for *CAVE Automatic Virtual Environment*) in operation at Defence R&D Canada – Valcartier. The CAVE, described later, displays a 3D model built from the received cumulative data from the perception, localization and attitude sensors of a robotic platform. The immersed operator uses a 3D input device to interact with the model and the robot avatar. Interactions range from changing the operator's vantage point, designating an area of interest from which the sensors will return data, defining waypoints or a path that the platform will follow. Motions and pose of the operator in the CAVE have a direct influence on visualization. A data exchange interface provides a bidirectional link between the *virtual world* in the CAVE and the platform, sensors and actuators in the *real world* (Fig. 2).

At the beginning of the project, it was decided to use a simulator instead of a real robot to quickly demonstrate the project's concept without having to deal with the issues related to the use of a real robot (error in position estimates, noisy scan data, spatial registration of sensor data, etc.). Therefore, the *real world* box in Fig. 2 was actually replaced with the *simulated world* box in Fig. 3.

Figure 3 describes the implementation of the system. The system is built around two main components: the *Immersive Environment* (the CAVE or virtual work cell) and the *Simulated World*. The former comprises the CAVE infrastructure and the associated systems required to capture the operator motions and intents while the latter contains the world in which the emulated robotic platform evolves, interacts and uses its sensors.

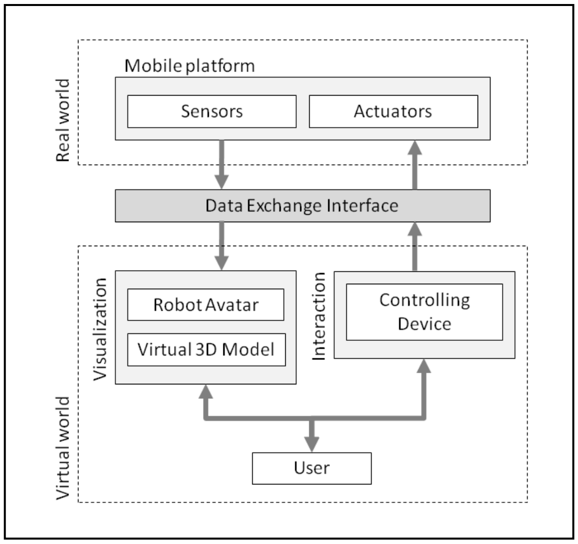


Figure 2. System concept description

The data exchange interface between the *Immersive Environment* and the *Simulated World* is based on a TCP/IP link. This interface supports different types of operations: *(1)* sending data captured by the robot's simulated sensors from the *Robot Manager* to the *Visualization Manager*; (2) sending the robot's pose from the *Robot Manager* to the *Visualization Manager*; and (3) sending control commands (such as move, rotate, scan) from the *Interaction Manager* to the *Robot Manager*.

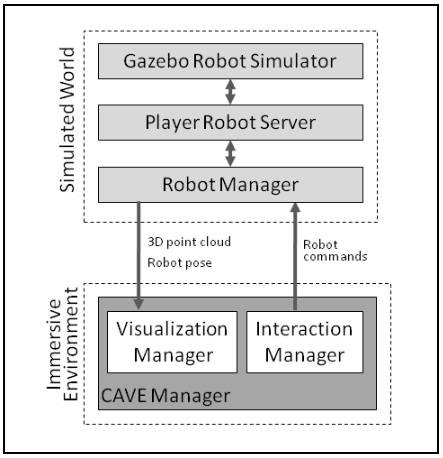


Figure 3. Implementation of the concept

IV. ROBOT SIMULATOR

The simulation part of the project comprises three components: the *Gazebo Robot Simulator*, the *Player Robot Server* and the *Robot Manager*, as shown in Fig. 3.

Gazebo is a robot simulator that is part of the Player project [11]. It allows simulating robots and their sensors / actuators and making those robots evolve in 3D virtual worlds with different types of objects. One of the advantages of this simulator is that it manages friction and collision detection.

The robot model used for the simulation in Gazebo is a Pioneer 2DX (from *ActivMedia*). A Hokuyo laser scanner is mounted on a panning unit on top of the Pioneer platform with the scanning plan perpendicular to the ground. With this setup the laser scanner can be rotated to obtain 3D volumetric scans of the environment such as the one shown in Fig. 4a. For the moment, the simulated laser scanner generates perfect readings but it is intended to apply an error model in the future. Different types of objects can be used to create the environment where the robot evolves. For our purposes, a virtual world of sufficient scale and complexity with walls and 3D geometric primitives (cube, sphere, cylinder and pyramid) was created.

The *Player Robot Server* provides the interfaces to interact with the simulator. Player is connected to both Gazebo and the *Robot Manager* and allows the *Robot Manager* to control the robot and to retrieve pose and laser scan data from the simulation.

The *Robot Manager* receives displacement and rotation commands (move forward, backward and turn right, left) from the *Interaction Manager* of the *Immersive Environment* and applies those commands to the simulated robot through Player. The *Robot Manager* also retrieves the pose of the robot along with the scans acquired by the laser scanner.

When scanning commands are received from the *Interaction Manager* of the *Immersive Environment*, the *Robot Manager* instructs the laser scanner to start scanning (perform full 360° scan). When the scan is completed, the *Robot Manager* converts the laser scan measurements from the laser scanner's coordinates frame (*range*, *phi*, *theta*) to the world coordinates frame (x, y, z) and sends them to the *Visualization Manager* of the *Immersive Environment*.

V. EXPLORATION BY IMMERSION

Immersive VR environments (such as CAVE) promote, among others, the ability to analyze and interpret a large amount of data faster and to explore this data in a more intuitive way (interact with visually displayed data). Immersive systems can also accelerate the understanding of complex data and enhance awareness.

The CAVE (from *Mechdyne*) at DRDC-Valcartier is a multi-person, four-sided, high-resolution 3D environment that is used for viewing and interacting with virtual content in an immersive interactive setting (Fig. 4a). To achieve realistic interactions with immersive displays the tracking system (*IS-900* system from *Intersense*) has fast update rates, low latency and smooth tracking. Consequently, it provides smooth and precise position and orientation (6-DOF) of the user's head while not interfering with the user's immersive experience. The tracking system offers ergonomically designed devices: a head tracker attached to stereo glasses (*CrystalEyes®3* from *StereoGraphics Corp.,* Fig. 4b) and a tracked wand (from *InterSense*) which incorporates buttons (four on one side and a trigger on the other side) and a joystick for interaction with the virtual content (Fig. 4c). Stereoscopic rendering based on stereo (active) glasses provides to the user the depth perception of virtual content.

The Human-centric immersive working environment is supported by a custom open architecture, mainly based on *openGL* and the 3D graphics toolkit *OpenSceneGraph*. For this project, two main modules have been developed: the *Visualization Manager* and the *Interaction Manager*.

A. *Visualization Manager*

In near real-time, the *Visualization Manager* displays on the four CAVE screens 3D data received from the *Robot Manager*. When a scan is performed by the laser scanner, sets of data (about 2000 data points per set, each point characterized by x, y, z and its relative height from the ground) are sent sequentially to the *Visualization Manager* over the TCP/IP communication link. The 3D point cloud is then displayed in the CAVE and each 3D point is represented by a color corresponding to its height (specific color spectrum is associated to height parameter). The virtual 3D model is then augmented with each received scan.

B. *Interaction Manager*

From the immersive VR environment, the user can control the robot by sending commands to it. The control and the navigation inside the immersive environment are executed by using the wand. Here are the commands and functionalities associated to the buttons on the wand:

a. *Blue* button: send a command to the robot to automatically scan a 360^{o} view of its environment;

b. *Green* button: send a command to the robot to start a manual scan of the area in front of the robot;

c. *Yellow* button: modify the camera mode: *(i)* "God view" (third person shooter), or *(ii)* robot view (first person shooter);

d. *Red* button: modify the control mode of the wand. Two modes have been defined: the *Robot Control* mode and the *Camera Control* mode. The *Robot Control* mode sends, via the wand joystick, commands to the robot to move forward, move backward, turn to the right, turn to the left, and also go to one specific position (using the wand trigger combined with the target visible on the screen). The *Camera Control* mode sends commands to translate the camera forward, backward, up, down, to the right, to the left, and also to rotate the camera toward the right / left around the z-axis (relative to the robot's position).

VI. CONCEPT DEMONSTRATION

A simulated *real world* was created in *Gazebo* with an instance of the Pioneer robot on which was installed a Hokuyo sensor mounted on a pan and tilt unit (Fig. 9a - the robot is located in the top right corner of the *real world* and Fig. 5 - a part of the simulated *real world* and the corresponding part in the virtual 3D model). This setup runs on one computer. The immersive applications (*CAVE Manager)* run on a separate computer.

The user wearing the stereo glasses stands in the center of the CAVE so as to be able to see the four screens where 3D graphics generated by the *CAVE Manager* are displayed. With these special glasses, the user can actually see objects in 3D and can walk around them, getting a proper view of what the objects look like (see Fig. 6 in which the user observes one scene from two different points of view at 180°, in standing and kneeling positions). When the user walks around in the CAVE or moves his head, his movements are tracked with the head tracker and the video adjusts accordingly. At this point, it is important to note that images in the figures are either taken from the user's point of view (no distortion) or out of the user's point of view (distortion is visible). The two images on Fig. 7 demonstrate distortion versus no distortion. The user uses the wand to control both the robot and the camera.

For the purpose of the demonstration only automatic scans were used. The user is presented with an empty virtual environment and only an avatar of the robot. The first thing the user does is to instruct the robot to take one full 360° scan in order to get an overview of where the robot is in the *real world* and to see what its surroundings are. This step starts the generation of the virtual 3D model (Fig. 4a).

Based on the partial representation of the environment obtained, the user selects an area to explore. The user then controls the robot (with the forward, backward left and right turns) to get to a new position (see Fig. 8a) where the next full 360° scan is to be performed. Fig. 8b shows the evolution of a partial 360° scan and Fig. 8c-e show the local refinement of two objects. These steps are repeated until the full *real world* is explored. Figure 9 shows the simulated *real world* and the final virtual 3D model obtained after the robot has completed the exploration.

VII. DISCUSSION AND FUTURE WORK

The work presented in this paper represents the initial step towards demonstrating the benefits of the use of immersive tools for remote robotic platform operation. There is still more work to be done to fulfill this objective.

Usability tests with the system have been performed with a limited number of users so far. Users who have tried the system were able to control the robot and navigate in the generated 3D model. They appreciated the fact that they had the possibility to visualize the explored portions of the environment which helped them to elaborate an efficient exploration path. Users also mentioned that being able to have a 3D representation of an object of interest would enhance their ability to decide on how to interact with this object.

It is clear that more extensive experiments involving target users will be required to assess the real benefits of the presented approach over traditional remote robotic operation using direct camera feeds. It is planned to have a groups of users exploring the same simulated environment using one of the two approaches. The results will be compared according to metrics such as time and efficiency and feedbacks from the users will be captured. Additional software will have to be developed to allow camera image acquisition in Gazebo and rendering of those images with the immersive display.

For the first phase of the project, it was decided to use a simulator to accelerate the development of a system prototype that could be used to perform experiments and that would serve as the base for further developments. The next phase of the project will require to transition from a simulated word to the real word. For this purpose, it is envisioned to use one of our Pioneer platforms equipped with a Hokuyo UTM-30LX sensor mounted on a pan and tilt unit. The transition to the real world will require that several challenges be addressed.

One of them is accurate positioning of the robot. In the simulation, it is assumed that the pose of the robot is known at any given time. For a real robot, an efficient positioning strategy has to be used. The intent is to implement a simultaneous localization and mapping (SLAM) strategy [12] based on odometry and iterative laser scans matching.

Another factor to take into account is the noise in the data generated by a real laser scanner. Combined with the positioning errors, noisy data introduce errors and artefacts in the generated 3D point cloud model of the environment. To mitigate this effect, it is envisioned to transition from a point cloud representation to a probabilistic modelling scheme such as the one presented in [13].

When the transition to a real world platform will be completed, it will be necessary to repeat the usability experiments in order to evaluate if the approach is efficient and viable when a real robot is used instead of a simulated one.

VIII. CONCLUSION

The ultimate goal of this project is to demonstrate the benefits of using immersive tools to explore an unknown environment with a remote mobile platform. In the first phase of the project, a prototype system that makes use of an immersive VR environment (CAVE) has been used to control a simulated robotic platform involved in an exploration task of an unknown environment. The immersive environment shows to the operator an incrementally built 3D model of the robot's area of operation. Initial experiments were performed where users navigated the robot and controlled its on-board sensor to explore and map a custom virtual world. Subsequent phases of the projects will involve more usability tests and the transition to a real world platform.

ACKNOWLEDGMENT

The authors would like to thank François-Michel De Rainville, Maxime Girardin and Frederic Drolet for their work on the project.

REFERENCES

[1] M.S. Kadavasal, and J.H. Oliver, “Sensor augmented virtual reality based teleoperation using mixed autonomy,” Journal of computing and information science in engineering, vol. 9, no. 1, 2009.

[2] J. Albiez, B. Giesler, J. Lellmann, J.M. Zöllner, and R. Dillmann, “Virtual immersion for tele-controlling a hexapod robot,” Climbing and Walking Robots, Springer Berlin Heidelberg, pp.1087–1094, 2006.

[3] N.Shiroma, R.Miyauchi, and F.Matsuno, “Mobile robot teleoperation through virtual robot,” Robot and Human Interactive Communication, pp.477-482, Aug. 2008.

[4] Q. Lin, and C. Kuo, “A virtual environment-based system for the navigation of underwater robots,” Virtual Reality, vol. 3, pp. 267-277, 1998.

[5] L. Yu, P.W. Tsui, Q. Zhou, and H. Hu, “A web-based telerobotic system for research and education at Essex,” Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatron., 2001.

[6] K.T. Simsarian, “A system for mobile robotic telepresence employing VR as the communication medium: interface metaphors,” Swedish Institute of Computer Science, Kista, 1999.

[7] S. Livatino, and F. Privitera, “3D visualization technologies for teleguided robots,” Proc. VR Software and Technologies, 2006.

[8] H. Martins, R. Ventura, “Immersive 3D teleoperation of a search and rescue robot using a head-mounted display,” Proc. of the 2009 IEEE Int. Conference on Emerging Technologies and Factory Automation (ETFA), Mallorca, Spain, Sep. 2009.

[9] F. Ferland, F. Pomerleau, C.T. Le Dinh, and F. Michaud. “Egocentric and exocentric teleoperation interface using real-time, 3D video projection”, Proc. of the ACM/IEEE International Conference on Human-Robot Interaction, pp 37–44, Mar. 2009.

[10] M. Koeda, Y. Matsumoto, and T. Ogasawara, “Development of an immersive teleoperating system for unmanned helicopter,” Proc of the 11th IEEE International Workshop on Robot and Human Interactive Communication, pp. 47- 52, 2002.

[11] B.P. Gerkey, R.T. Vaughan, and A. Howard. “The Player/Stage Project: Tools for Multi-Robot and Distributed Sensor Systems,” Proc. of the Intl. Conf. on Advanced Robotics (ICAR), pp.317–323, Coimbra, Portugal, June 2003.

[12] S. Thrun, “Simultaneous localization and mapping,” Spatial Mapping Approaches in Robotic and Natural Mapping Systems, Springer, 2006.

[13] J. Fournier, B. Ricard, and D. Laurendeau, “Mapping and exploration of complex environments using persistent 3D model,” Proc of the Fourth Canadian Conf. on Computer and Robot Vision, pp. 403–410, 2007.

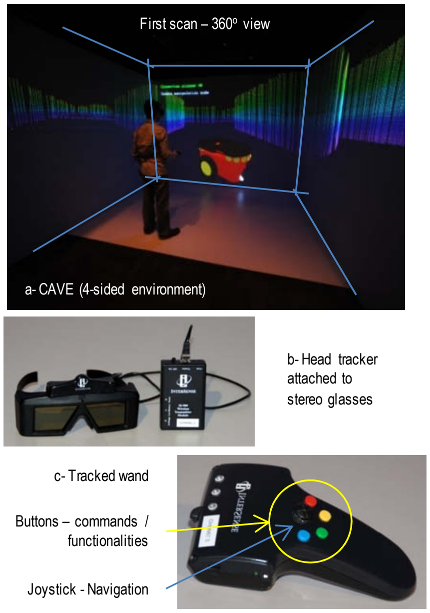


Figure 4. Immersive visualization system

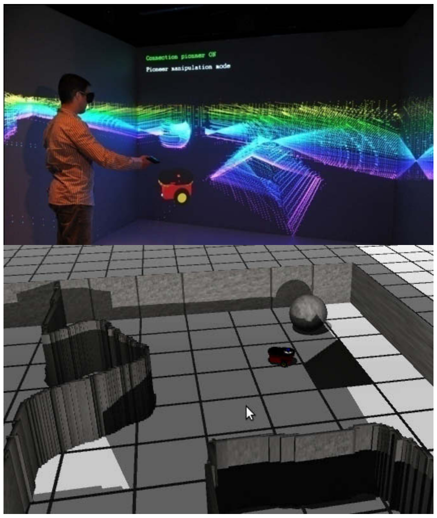


Figure 5. Simulated robot in the simulated *real world* (*bottom*) and robot’s avatar in the immersive VR environment (*top*)

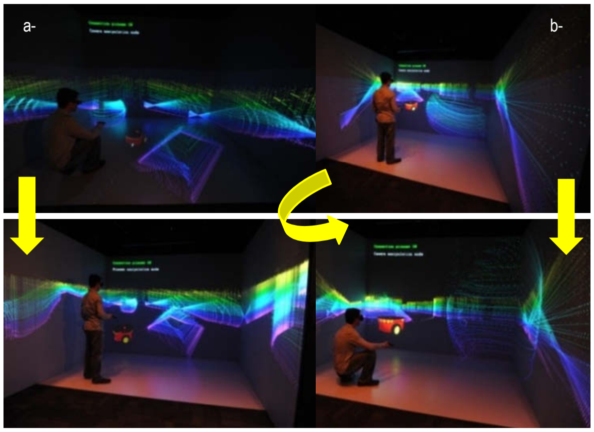


Figure 6. 3D model observation : one scene viewed from two points of view (a- and b-)

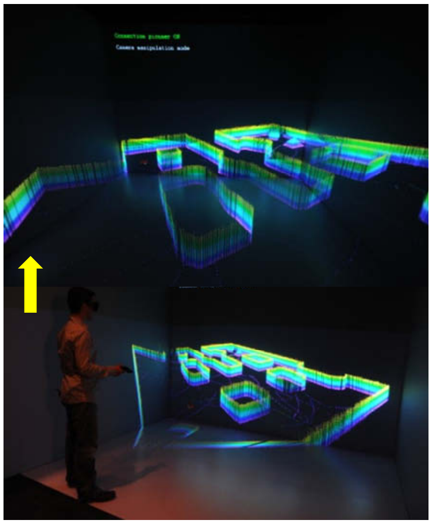
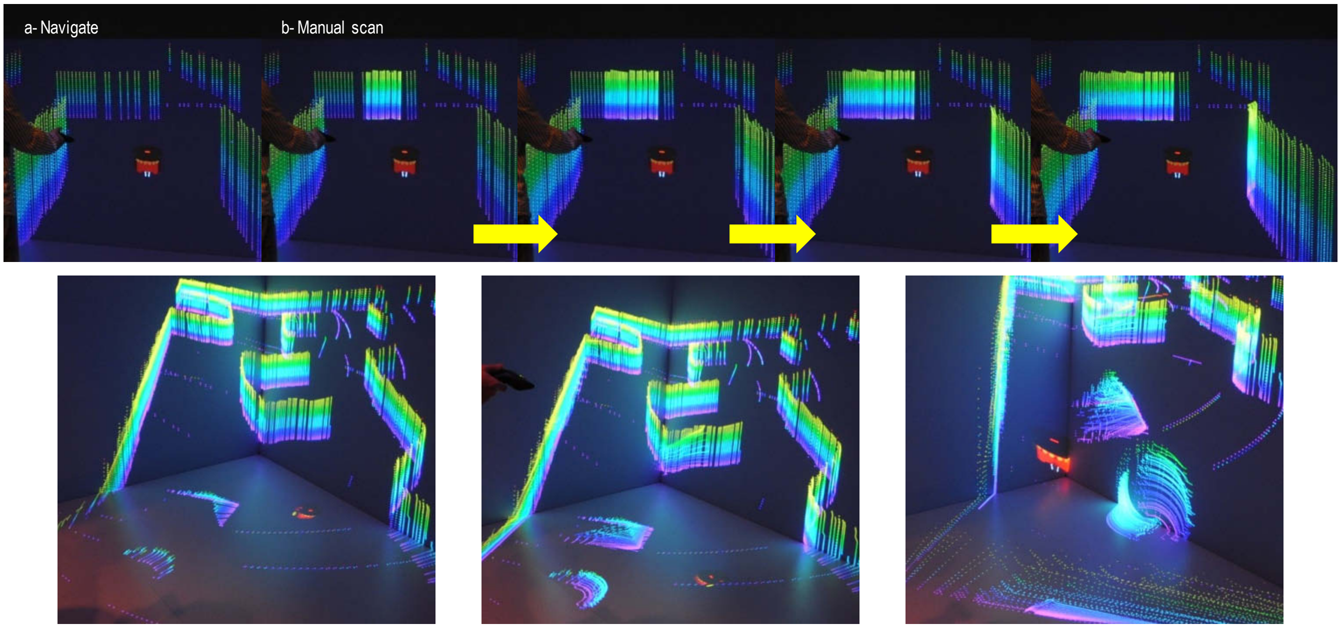


Figure 7. From “God view” (*bottom*) to “real size” (*top*); Image distortion understanding – 3D model viewed from the user’s perspective (*top*) versus a perspective out of the user (*bottom*)



c- Move around one “object”

d- Manual scan – Refine “object”

e- Move, rotate and scan

Figure 8. Robot control and 3D model building: a- Navigate to a specific area; b- Manual scan, the robot stays at the same position but the panning unit rotates on itself; c- to e- Local refinement

a- Simulated *real world* in which the Pioneer-like robot evolves

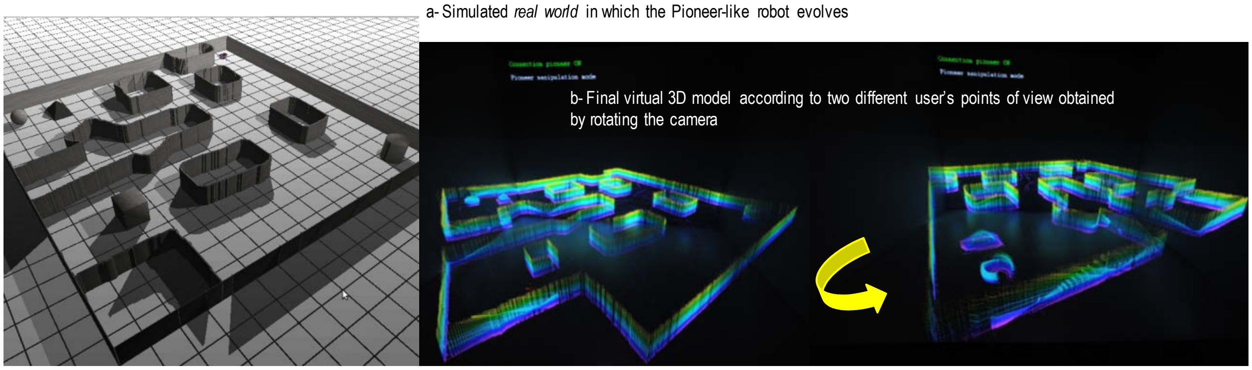


Figure 9. From the simulated *real world* to the 3D virtual model