# Remote Driving With a Multisensor User Interface

**Gregoire Terrien**
Institut de Systèmes Robotiques, L'Ecole Polytechnique Fédérale de Lausanne

**Terrence Fong and Charles Thorpe**
The Robotics Institute, Carnegie Mellon Univ.

**Charles Baur**
Institut de Systèmes Robotiques, L'Ecole Polytechnique Fédérale de Lausanne

## ABSTRACT

Remote driving is a difficult task, primarily because operators have problems understanding the remote environment and making control decisions. To make remote driving easier and more productive, we are developing sensor fusion techniques using range sensors to build active, sensor fusion based interfaces. In our work, we use sensor fusion to facilitate human perception and to enable efficient command generation. In this paper, we describe a multisensor user interface for remote driving.

## INTRODUCTION

Perhaps the most difficult aspect of remote driving is that the operator is unable to directly perceive the remote environment. Instead, he is forced to rely on sensors, bandwidth-limited communications links, and an interface to provide him with information. As a result, the operator often fails to understand the remote environment and makes judgements errors. Thus, to make remote driving easier, we need to find ways to facilitate information transfer and to improve situational awareness.

Our approach is to develop sensor fusion based operator interfaces. Sensor fusion has long been used to improve automatic processes such as mapping. However, we believe that sensor fusion can also be used to create capable and compelling interfaces. In particular, we are using sensor fusion to: (1) create displays which enable better understanding of the remote environment; and (2) efficiently and accurately generate motion commands.

In [1], we described a system for fusing and displaying stereo vision and sonar data. In this paper, we describe extensions to our previous work. Specifically, we discuss the addition of a scanning laser rangefinder (ladar), integration of the multisensor platform with a mobile robot, and the development of command generation tools.

## RELATED RESEARCH

### SENSOR FUSION DISPLAYS

Sensor fusion displays combine information from multiple sensors or data sources to present a single, integrated view. Sensor fusion displays are important for applications in which the operator must rapidly process large amounts of multi-spectral or dynamically changing heterogeneous data.

In military aerospace, sensor fusion displays combine information from imaging sensors (visible light cameras, night-vision devices, millimeter wave radar, thermal imagers, etc.) and databases (digital maps, target catalogs). The resultant displays are used to improve cockpit efficiency during target acquisition or tracking, tasks which demand high-levels of situation awareness and cognitive decision making[6].

In civil air transport, sensor fusion displays are being considered for use in enhanced or synthetic vision systems. These displays would enable pilots to better detect runway features and incursions during landing, and would aid in detecting obstacles and traffic in taxi[5]. Sensor fusion displays would also enable airport traffic controllers to operate in low-visibility weather conditions, i.e., the sensors would allow controllers to "see" aircraft movements through fog or cloud.[4].

More recently, sensor fusion displays have been used as control interfaces for telerobots. In particular, the Virtual Environment Vehicle Interface (VEVI) combines data from a variety of sensors (stereo video, ladar, GPS, inclinometers, etc.) to create an interactive, graphical 3D representation of the robot and its environment. For example, multiple types of range data are used to construct polygonal models of explored terrain[3].

## AUGMENTED REALITY

Augmented reality is a variation of Virtual Environments (VE), otherwise known as Virtual Reality. Whereas VE's completely replace reality by immersing the user inside a synthetic environment, augmented reality allows the user to see the real world (often using a head-mounted, see-through display) in combination with information superimposed or composite on the display. Thus, augmented reality enhances a user's perception of and experience with the real world[8].

Augmented reality has been used to assist users in outdoor navigation tasks. In [11], a wearable computer system and see-through display provide location-specific multimedia information to Columbia University visitors. Differential GPS (dGPS) and a magnetometer/inclinometer are used to track a user's location and to update the display with information such as current location, nearby points of interest and point-to-point directions. Similarly, in [9], dGPS and a digital compass are used to create graphical overlays in a see-through display to assist users in large-area terrestrial navigation.

Augmented reality has also been used to assist remote collaboration. In [10], a collaborative system provides a local user with direct, unmediated access to the output of sensors (proximity, location, electronic tags) attached to a remote user's wearable computer. This allows the local user to provide context-specific information to the remote user. For example, an expert is able to provide a field technician with supplemental information (schematics, specifications, etc.) based on the sensed location.

## SYSTEM CONFIGURATION

### PREVIOUS WORK

We previously developed a multisensor system incorporating stereo vision, an ultrasonic sonar ring, and odometry[1][2]. We chose these range sensors based on their complementary characteristics: stereo vision provides passive wide-area ranging with high angular resolution but with non-linear depth measurements; sonar provides active sector-based ranging with linear depth measurements but with low angular resolution.

We used the multisensor system to create two sensor fusion displays: a 2D image with color overlay and a local map. The 2D image showed range data as colors overlaid on a monochrome 2D camera image. We found that this display helped users make relative distance judgements and to spot nearby obstacles. The local map was constructed via histogramic filtering of the range data into an occupancy grid. With the high angular resolution from stereo, object contours were clearly visible and the map improved users' situational awareness.

The system, however, had two significant weaknesses. First, we found that the sonar/stereo combination failed in certain environments. For example, smooth surfaces with low texture (e.g., a white wall) were frequently missed by both the sonars (specular reflection) and the stereo (poor correlation). Second, the sensors were mounted on an electric wheelchair equipped with wheel encoders. While this enabled us to show that sensor fusion displays improved understanding of the remote environment, it did not allow us to use the displays for command generation (i.e., users were unable to teleoperate the wheelchair).

## LADAR INTEGRATION

To address our system's sensing inadequacies, we added a Sick "Proximity Laser Scanner" (PLS) ladar to the sensor suite[7]. Ladar sensors provide precise range measurement with very high angular resolution, but are usually limited to a narrow horizontal band (i.e., a half-plane). This forms a good complement to the sonar and stereo sensors, which are less accurate but have a broader field-of-view. The PLS ladar has 5 cm accuracy over a wide range (20cm to 50m), a 180 degree horizontal field- of-view (361 discrete measurements) and greater than 5 Hz scan rate.

We should note that ladar is not, by itself, a panacea. Although this sensor provides high resolution ranging independent of surface type and is well suited for mapping certain environments (e.g., indoor office walls), it also has limitations. Since the ladar measures depth in a narrow horizontal band, obstacles which do not intersect the scanning plane will be missed. Moreover, obstacles with varying vertical profiles (e.g., a table) will not be correctly measured (e.g., only the table legs may be scanned). Additionally, both smoke and steam reduce beam intensity, thus producing erroneously large range measurements (due to weak return) or complete failure (when the reflected signal is too weak to be detected). Finally, ladar is susceptible to glare which often makes it problematic for outdoor use.

Figure 1 shows our current multisensor system. The ultrasonic sonar ring uses Polaroid 600 series electrostatic transducers and provides time-of-flight range at 25Hz. The stereo vision system is a Small Vision Module[12] and produces 2D intensity (monochrome) images and 3D range (disparity) images at 5 Hz. Odometry is obtained from wheel-mounted optical encoders.
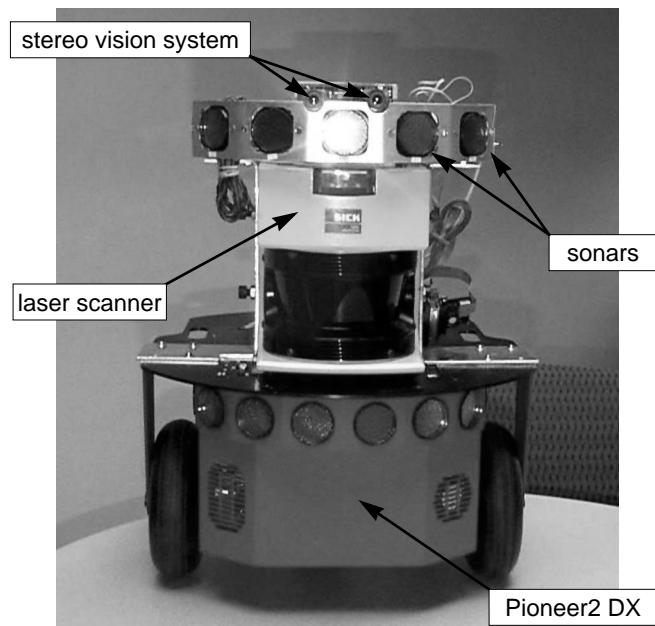
Figure 1. Multisensor system

Table 1 lists situations commonly encountered in indoor vehicle teleoperation. Although no individual sensor works in all situations, the collection of sensors provides complete coverage.

Table 1. Sensor performance in teleoperation situations

| Situation | 2D Image (intensity) | 3D Image (disparity) | Sonar | Ladar (laser) |
|---|---|---|---|---|
| smooth surfaces (no visual texture) | OK | Fails[a] | Fails[b] | OK |
| rough surface (little/no texture) | OK | Fails[a] | OK | OK |
| far obstacle (> 10m) | Fails[c] | Fails[d] | Fails[e] | OK |
| close obstacle (< 0.5m) | OK[f] | Fails[g] | OK[h] | OK[i] |
| small obstacle (on the ground) | Fails[c] | OK | OK | Fails[j] |
| dark environment (no ambient light) | Fails | Fails | OK | OK |

a. no correlation
b. specular reflection
c. no depth measurement
d. poor resolution
e. echo not received
f. limited by focal length
g. high disparity
h. limited by transceiver
i. limited by receiver
j. outside of scan plane

## ROBOT INTEGRATION

To investigate remote driving, we mounted the sensors on a Pioneer2[1] DX (P2DX) mobile robot (Figure 1). The P2DX has differential drive and is designed for indoor environments. We equipped the P2DX with on-board computing (Ampro P5e) and wireless ethernet (Lucent

---

1. Pioneer is a trademark of ActivMedia, Inc.

WaveLan). On-board computing is used for sensor management (data collection and transmission), obstacle avoidance, and position servoing.

The obstacle avoidance algorithm differs from conventional methods because it does not perturb the robot's trajectory near obstacles. Instead, the algorithm scans for obstacles in the direction of motion and slows or stops the robot's translation and rotation rates[7]. This allows the robot to avoid collisions without sudden directional changes which would confuse the operator.

## USER INTERFACE

We developed a remote driving interface which contains sensor fusion displays and a variety of command generation tools. The interface is designed to improve situational awareness, facilitate depth judgement, support decision making and speed command generation. In particular, we placed considerable emphasis on creating effective affordances and representations so that data is readily accessible and understandable. Additionally, since the operator has to generate remote driving commands by analyzing what is displayed on the screen, we tried to provide an interface which is intuitive, coherent, and maximizes information transfer.
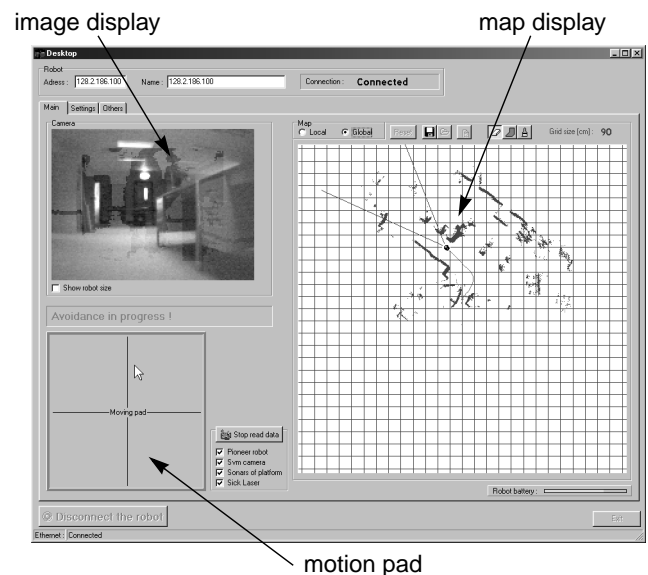


Figure 2. Sensor fusion user interface for remote driving

Figure 2 shows the main window of our sensor fusion based user interface. The interface contains three primary tools: the image display, the motion pad, and the map display. In addition, to enable the operator to better understand the remote environment and to better make decisions, we developed tools for measuring distance, checking clearance, and for finding correspondences between map and image points.

## Image Display

The image display contains a monochrome video image with a color overlay to improve depth judgement and obstacle/hazard detection. Hue values encode depth information from close (yellow) to far (blue). Since close depth is more relevant (e.g., for identifying and avoiding nearby obstacles), we vary hue exponentially (i.e., near ranges are encoded with more values than far ranges).

## Motion Pad

The motion pad enables the operator to directly control the robot. Clicking on the vertical axis commands a forward/reverse translation rate. Clicking on the horizontal axis commands a rotation rate. Translation and rotation are independent, thus the operator can simultaneously control both by clicking off-axis. The pad's border color indicates the robot's status (moving, stopped, etc.).

## Map Display

To navigate the robot, we created a map display which gives the operator with a bird's eye view of the remote environment. The display is constructed as the robot moves and shows sensed environment features and the robot's path.

The map display provides a local and a global map. With the local map, the user can precisely navigate through complex spaces. For large-area navigation, the global map helps maintain situational awareness by showing where the robot has been.

At any time, the operator can annotate the global map by adding comments or drawing "virtual" obstacles. For example, if the operator finds something of interest, he can label the map. Additionally, if he decides that a particular region is dangerous, he can draw an artificial barrier on the map and the robot's obstacle avoidance will keep the robot from entering the region (see Figure 8).

## MODULES AND DATA FLOW

The system architecture is shown in Figure 3. The robot is driven by rate or pose commands generated by the interface. Pose commands are processed by a path servo which generates a smooth trajectory from the current position to the target pose. All motion commands are constrained by the obstacle avoidance module.

All sensors are continuously read on-board the robot and the data transmitted to the interface. The sensor readings are used to update the image and map displays. Fusion algorithms for both displays are described in the following sections.
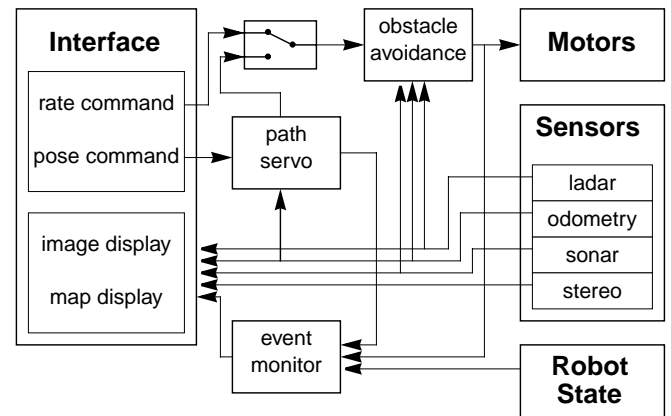
Figure 3. System architecture

An event monitor watches for critical system events and mode changes (e.g., obstacle avoidance in progress). It also continually monitors robot health and generates appropriate status messages to be displayed to the user.

## SENSOR FUSION ALGORITHMS

## Image Display

We create the image display by overlaying range information as colors on a monochrome image taken from one of the stereo cameras. This method does not provide an absolute indication of range. However, we find it greatly improves relative distance judgement.
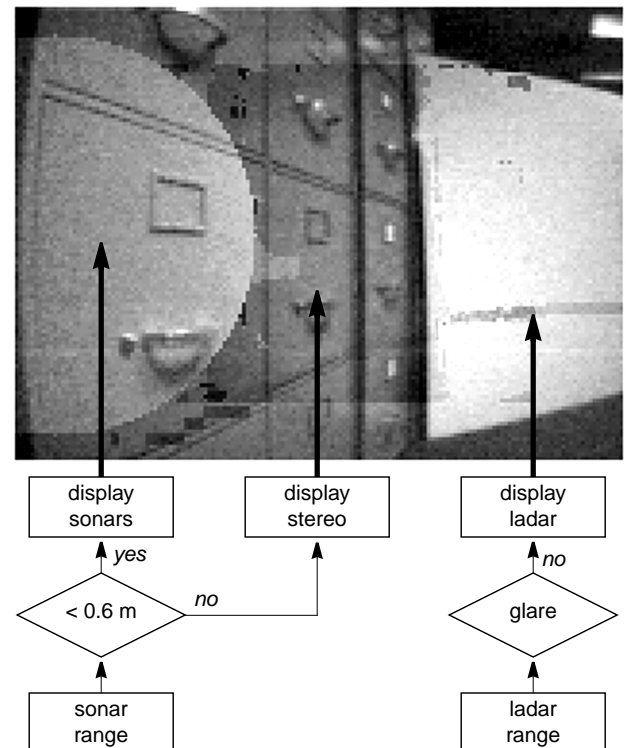
Figure 4. Image display processing

Figure 4 shows how the image display is constructed. For each overlay pixel, the sonar range is used to decide whether to display sonar or stereo data. When the sonar

range is low, sonar data is used because stereo correlation fails when objects are too close. Otherwise, if the sonar range is high, stereo is displayed. In addition, because the ladar is precise and reliable in an office environment, we always overlay ladar range when available (i.e., unless the ladar detects glare).

<u>Map Display</u>

We create the map display using sensor data and vehicle odometry for registration. We currently do not perform any filtering or adaptive registration of sensor data. Thus, map accuracy is highly correlated with odometry errors. The interface allows the user to select which sensors are used for map building at any time.
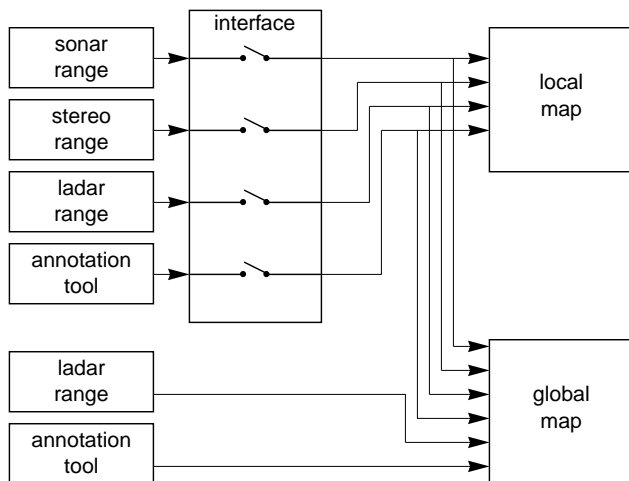


Figure 5. Map display processing

Figure 5 shows how the map is constructed. The local map only shows current sensor data in proximity to the robot. Past sensor readings are discarded whenever new data is available. In contrast, the global map displays sensor data over a wide area and never discards sensor data. Additionally, the global map allows the user to add annotations.

## RESULTS

### IMAGE DISPLAY

To evaluate the image display, we placed the robot in a setting which has difficult to sense characteristics: in front of the robot is a smooth, untextured wall; close to the robot is a large office plant. Figure 6 shows the image display for this scene with various overlays. As the figure shows, each range sensor individually has problems, but collectively provides robust sensing of the environment.

In the top left image (stereo only), the wall edges are clearly detected and the plant partially detected (the left side is too close for stereo correlation). However, the cen-



*Stereo only*      *Sonar only*

*Sonar and stereo*      *Ladar only*
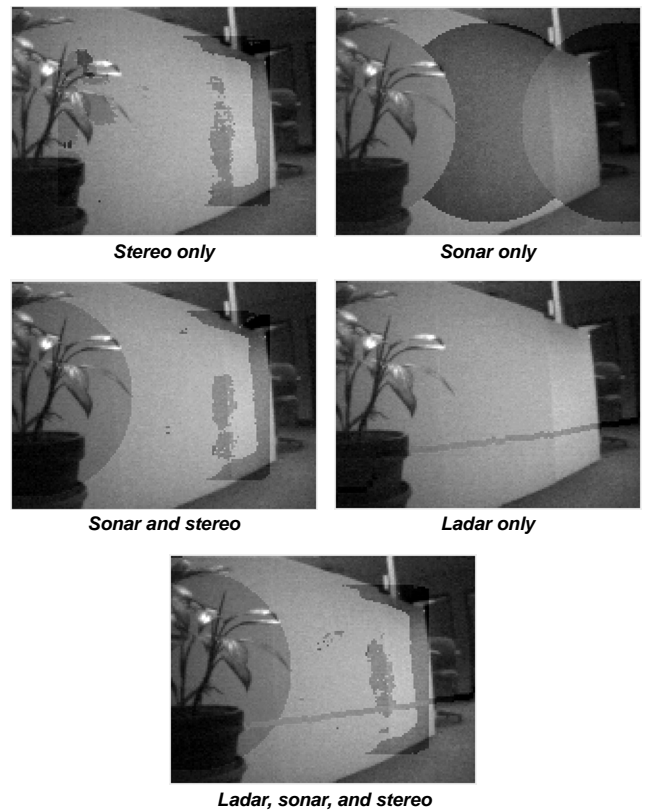
*Ladar, sonar, and stereo*

Figure 6. Sensor fusion based image display

ter of the wall (untextured) is completely missed. In the top right image (sonar only), the plant is detected well, but the wall is shown at incorrect depths due to specular reflection. In the middle left image (fused sonar and stereo), both the wall edge and plant are detected, but the center remains undetected. In the middle right image (ladar only), we see that the wall is well defined, but that the planar scan fails to see the plant. In the bottom image (all sensors), we see that all features are properly detected. The sonars detect the plant, the ladar follows the wall, and stereo finds the wall edge.

## MAP DISPLAY

<u>Map Building</u>

To evaluate map building, we placed the robot in a room with a variety of surfaces (smooth, rough, textured, non-textured). Figure 7 shows maps constructed with different sensors combinations.

In the first image (stereo only) we see some clearly defined corners, but some walls are not well detected due to lack of texture. In the second image (sonar only), the sonar's low angular resolution and specular reflections result in poorly defined contours. In the third image (stereo and sonar), both corners and walls are well detected, however due to stereo's non-linear depth accuracy, there is significant error. In the final image (ladar only), the map clearly shows the room. Obviously, for an

**Stereo only**     **Sonar only**

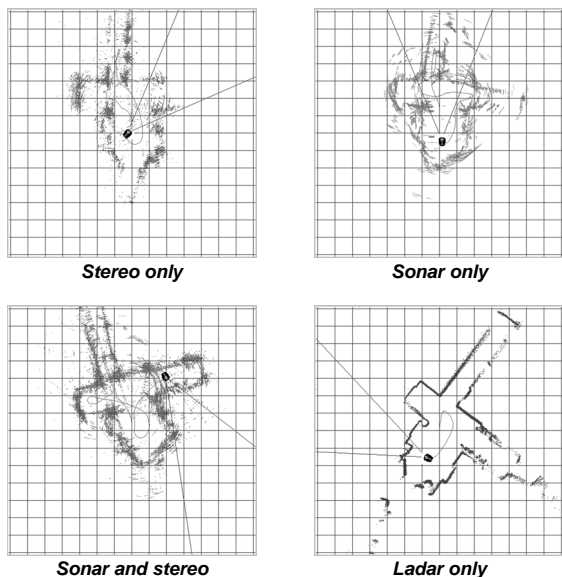**Sonar and stereo**     **Ladar only**

Figure 7. Map display

indoor environment in which the principal features are uniformly vertical (i.e., walls), the ladar produces the most useful maps.

### Map Annotation

We found that annotation greatly improved the usefulness of the map display in two ways. First, labeling helped operators to preserve contextual information. Second, "virtual obstacles" enabled operators to add critical information which was missed by the sensors.

Since the maps only show a 2D projection of the world, we found that obstacles with irregular profiles are difficult to distinguish without labeling. For example, chairs are often hard to identify: they may be shown as a set of points (legs sensed), a line (horizontal edges sensed), or a scattered cloud of points. By allowing the user to add labels, features on the maps became easier to interpret.

We found that by letting operators create "virtual" obstacles, the safety and robustness of remote driving was improved. For example, by drawing lines on the map, operators were easily able to keep the robot away from dangerous, but difficult to sense obstacles.
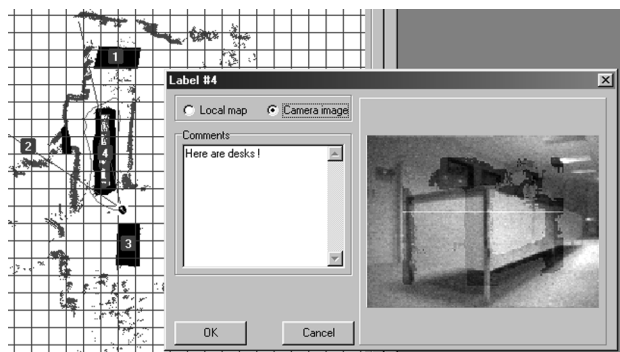


Figure 8. An annotated map

Figure 8 shows an annotated map with several labels (numbered 1 to 4). As the figure shows, it is not evident that the center of the map (label 4) shows a set of desks. Thus, the operator has chosen to mark this region with the comment "Here are desks" and to include a camera image. Figure 8 also shows "virtual" obstacles (black lines and regions): a door (near label 2) and unsafe areas (table: labels 1 and 3) have all been fenced off.

## OBSTACLE DETECTION

One of the most challenging tasks during vehicle teleoperation is obstacle detection. Although no sensor exists which always detects all obstacles, by exploiting complementary sensor characteristics, we can avoid individual sensor failures and improve obstacle detection.
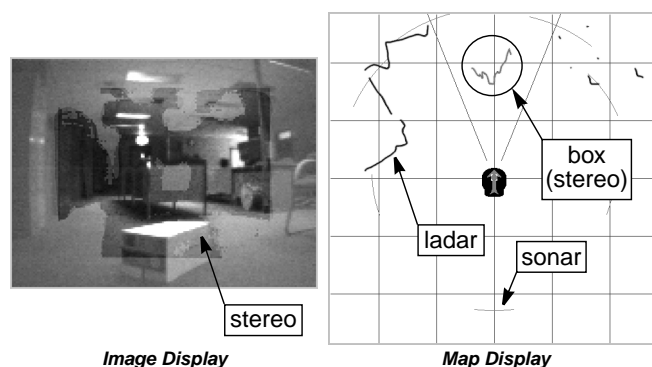


**Image Display**     **Map Display**

Figure 9. Detection of small obstacle

Figure 9 shows a scene with a box on the floor. Because the box is too small, it is not detected by the ladar (it is too short to intersect the scanning plane) nor by the sonars (it is located outside the sonar cones). However, it is properly detected by stereo as both displays show.
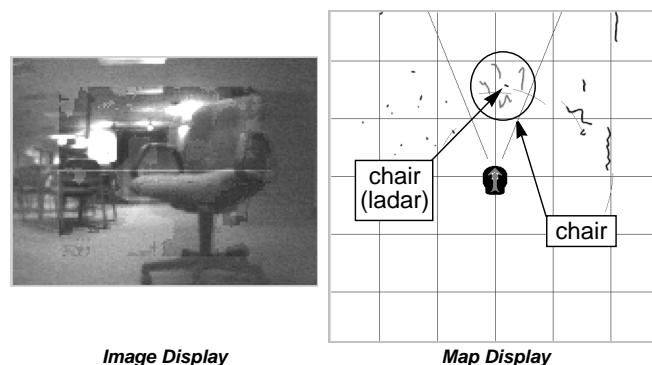


**Image Display**     **Map Display**

Figure 10. Detection of a chair

Figure 10 shows a situation in which the robot is approaching a chair. We can see that the chair is well detected by the stereo camera and the sonars. However, the ladar has problems with the chair because only the supporting post intersects the scanning plane (resulting in the a tiny spot on the map).

## COMMAND GENERATION

In order for remote driving to be efficient, the interface must provide capable command generation tools. Thus, in addition to the motion pad, we used the image and map displays to generate position commands. Double-clicking on the image selects a target point: heading is computed by the offset from the central image axis and distance is based on range data. Similarly, double-clicking on the map selects a relative position. In both cases, once a target point is designated, the robot computes a path and begins moving to that position.

If the robot discovers an obstacle during the automatic motion, it automatically slows down and informs the operator (see Figure 2). If the obstacle becomes too close, the robot stops and warns the operator that it cannot continue in this direction. At any time during the automatic motion, the user can again take control of the robot by clicking on the motion pad.

## DECISION MAKING TOOLS

### Distance Measurement

Most human senses are relative and not absolute. For example, we can say if something is hot or very hot, but not what is the absolute temperature. Human vision works much the same way: judging relative distance is easier (and more precise) than estimating absolute distances. For remote driving, however, distance information is indispensable for choosing paths and making decisions. Thus, we added a measurement tool to the image display (Figure 11).
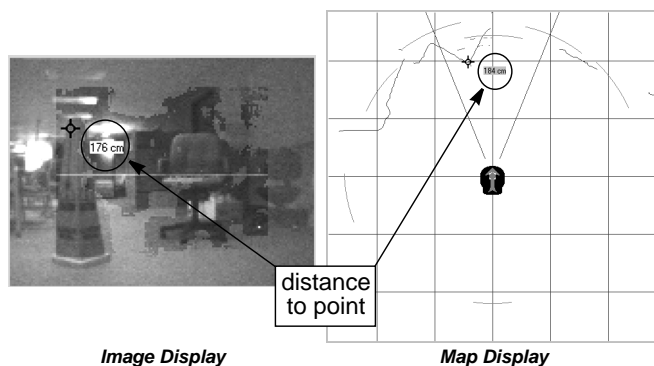


Figure 11. Depth measurement

When the operator clicks on the image or the map display, the interface automatically computes the absolute distance based on the available range data. The distance is then overlaid on the display and the corresponding point is highlighted (e.g., if the user clicks a point on the image display, the corresponding point is shown on the map display).

### Clearance Check

Another significant problem faced during remote driving is clearance estimation: will the vehicle be able to pass through an area or underneath something without getting stuck? Thus, we provided a tool for checking clearance.
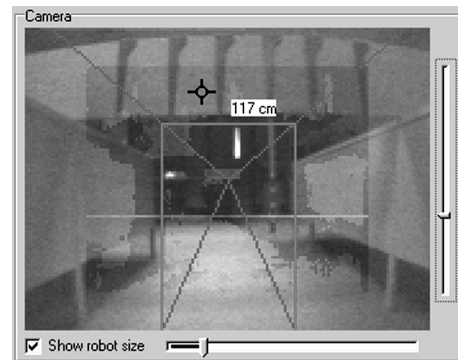


Figure 12. Clearance check

The robot's dimensions are represented by a rectangle overlaid on the image (Figure 12). The size of the rectangle is scaled based on the point being checked. Also, the rectangle is colored to reflect depth. When the operator clicks anywhere in the image, the tool projects the rectangle and displays the clearance required by the robot.

### Finding Correspondences

In some situations, the operator has difficulties finding the correspondence between a point on the map and a point on the image (and vice-versa). Since this information is critical for navigating and for accurately annotating the map, we implemented a tool to display matching points (Figure 13).
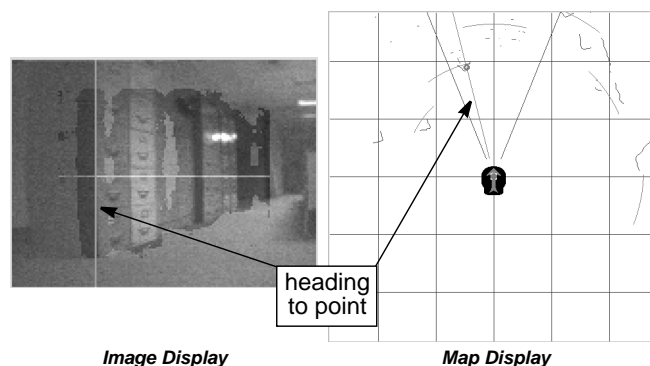


Figure 13. Finding correspondences

When the operator clicks on the image, the distance to the point and the corresponding map point (with the vector to the point) are shown. If the depth of the chosen point is not available, only the vector is drawn. Similarly, when the operator clicks on the map, a vertical line is drawn on the image: a point on the map corresponds to a vertical line in the image since the map is an orthogonal projection. The line is colored based on distance.

## FUTURE WORK

A number of improvements would make our system more reliable. First, we currently estimate robot position with dead-reckoned odometry. Since we use robot position to register sensor data, better positioning would improve accuracy. Second, our map building algorithm is very simplistic: we perform no data filtering and no registration. However, by applying techniques such as[13], we could create large-area maps and benefit from map-based localization. Finally, a higher level of robot autonomy would lessen the impact of operator differences and handicaps on system performance. For example, improving the robot's ability to identify hazards would reduce the need for the operator to be in continuous control.

To date, we have only collected anecdotal evidence that the sensor fusion interface improves remote driving. In order to rigorously assess usability and to better understand which features/affordances are most helpful, we need to conduct formal user studies. In particular, we would like to quantify how each display and tool improves or hinders task performance. Moreover, we need to understand the interface's limitations: when does it work well and when does it fail?

Lastly, we have carefully tuned our system for remote indoor driving based on sensor characteristics and environmental constraints. For example, our fusion algorithms give highest priority to ladar data because the ladar is ideal for sensing walls and objects in an office environment. But, if we wish to perform a different task, such as exploration in natural terrain, in which the ladar would be less effective (due to smoke/fog effects, distance to obstacles, etc.), we need to understand how to modify the system (e.g., what weighting factors need to be changed).

## CONCLUSION

The guiding principal in our work is that intelligent interfaces between humans and robots improve teleoperation performance. In particular, we believe that truly integrated and efficient human-robot systems can only be achieved through the development of better interfaces.

We have found that with an appropriate sensor suite and user interface, sensor fusion is a powerful method for improving vehicle teleoperation. In our system, we used a suite of sonar, stereo vision, and ladar sensors to remotely drive a mobile robot. Through our work, we demonstrated that a multisensor interface makes it easier to understand the remote environment, to assess the situation, to make decisions and to effect control.

## REFERENCES

1. Meier, R., Fong, T., Thorpe, C., and Baur, C., "A Sensor Fusion Based User Interface for Vehicle Teleoperation", Field and Service Robots, Pittsburgh, PA, August 1999.

2. Meier, R., "Sensor Fusion for Teleoperation of a Mobile Robot", Diplome Thesis, Swiss Federal Institute of Technology Lausanne (EPFL), March 1999.

3. Hine, B., et. al., "VEVI: A Virtual Environment Teleoperation Interface for Planetary Exploration", SAE 25th ICES, San Diego, CA, July 1995.

4. Foyle, D., "Sensor Fusion Display Evaluation using Information Integration Models in Enhanced/Synthetic Vision Applications", AVID Workshop, NASA Conference Proceedings 10128, Moffett Field, CA, December 1993.

5. Foyle, D.C., et. al., "Enhanced/synthetic vision systems: Human factors research and implications for future systems", Paper 921968, SAE Aerotech Meeting, 1992.

6. Wise, J., "Design of Sensor Fusion Displays: Human Factors and Display System Guidelines". Westinghouse R&D Center Research Report 87-1C60-SCVIS-R1, 1987.

7. Terrien, G., "Sensor Fusion Interface for Teleoperation", Diplome Thesis, Swiss Federal Institute of Technology Lausanne (EPFL), March 2000.

8. Azuma, R., "A Survey of Augmented Reality", in Presence: Teleoperators and Virtual Environments 6(4), August 1997.

9. Thomas, B., et. al., "A Wearable Computer System with Augmented Reality to Support Terrestrial Navigation", Second International Symposium on Wearable Computers, Pittsburgh, PA, October 1998.

10. Bauer, M., et. al., "A Collaborative Wearable System with Remote Sensing", Second International Symposium on Wearable Computers, Pittsburgh, PA, October 1998.

11. Hollerer, T., Feiner, S., and Pavlik, "Situated Documentaries: Embedding Multimedia Presentations in the Real World", Third International Symposium on Wearable Computers, San Francisco, CA, October 1999.

12. Konolige, K., "Small Vision System: Hardware and Implementation", Eight International Symposium on Robotics Research, Hayama, Japan, 1997.

13. Thrun, S., et. al., "A Real-Time Algorithm for Mobile Robot Mapping with Applications to Multi-Robot and 3D Mapping", IEEE ICRA, San Francisco, CA, April 2000.