

## An Advanced Teleoperation Testbed

Bill Ross<sup>1</sup>, John Bares<sup>1</sup>, David Stager<sup>1</sup>, Larry Jackel<sup>2</sup>, Mike Perschbacher<sup>3</sup>

1. The Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA. br@cs.cmu.edu
2. DARPA TTO Project Manager
3. RovnoTech

**Abstract:** *Due to the technology available, most previous work in teleoperated robotics used relatively low-resolution video links and provided limited perceptual feedback to the teleoperator. In most cases, these projects reported only limited teleoperator success compared to vehicles with human drivers on-board. We set out to build a high-fidelity teleoperation system which takes advantage of recent technological advances. This system permits highly capable teleoperation and has allowed us to begin to investigate the minimum system requirements for effective teleoperation.*

### 1. Introduction

While there have been a great number of teleoperated machines built, including some practical machines in daily commercial use, teleoperation success in highly unstructured environments has been comparatively limited. The reason for this may be simply that high resolution video cameras, video transmission systems, and displays have only recently become readily available.

The goal of this project was to develop a very high-fidelity teleoperation system for the CMU Crusher robot, to test this system in challenging, real-world terrain, and to produce estimates of the minimum parameters of an highly effective teleoperation system for use in unstructured environments.

Sponsored by Defense Advanced Research Projects Agency TTO Program under DARPA CMO Agreement MDA972-01-9-0005. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressly or implied, of the Defense Advanced Research Projects Agency or the U.S. Government. Approved for Public Release, Distribution Unlimited.

## 2. Background

There have been a number of excellent studies done on the impact of system capabilities on a teleoperator's performance [1-10]. In most of these studies the conclusion was that teleoperators seldom perform as well as human drivers and often performed poorly, with collisions and vehicle rollovers being common [10]. For the most part, however, these studies tested only relatively low-quality video systems with limited resolution and field of view, and with high latencies. Most of these tests were also conducted either in simulation[1][5], or in real-world terrain with relatively structured paths providing only limited challenges for teleoperators.

## 3. System Overview

While the performance targets for our system were ambitious, we never contemplated building a permanent system for continuous use. In many cases, we used what was available quickly and affordably to get the system up and running.



Figure 1: The Crusher robot

*Vehicle:* The system is based on the UGV Crusher robot. Crusher is a 6,800 kg, six-wheeled, hybrid-powered robot that has shown to be extremely capable in the most severe terrain. Crusher's active suspension provides a smooth ride for the camera system over rough ground at up to 12 meters per second.

*Cameras and displays:* The system is based on five 1600 by 1200 pixel, single-CCD, Bayer-pattern color cameras which give an aggregate field of view of 200 by 30 degrees. The overall resolution is 40 pixels per degree (1.5 MOA per pixel) in the field of view. This resolution is about 4 times the linear pixel resolution provided by typical television systems. Images are displayed on standard

LCD monitors with matching resolution which are sized and positioned so that the operator's field of view closely matches that of the cameras in the real world.

*Communications:* Primary communications between the vehicle and the control station is carried by a 1 km single-mode CWDM optical fiber which can handle 16 separate data channels. Currently we use five of these channels for the video feeds and a sixth for a gigabit Ethernet connection to the vehicle which carries vehicle IMU data, vehicle audio, and control information. The total data transmitted over the fiber is about 8.5 gigabits per second with almost no latency added by the fiber link. The fiber itself is a 1.6mm high-strength tactical cable with an aramid strength member and a polyurethane jacket which has proven to be surprisingly robust. The fiber is automatically reeled in and out using an iRobot spooler which is slaved to the vehicle odometry.

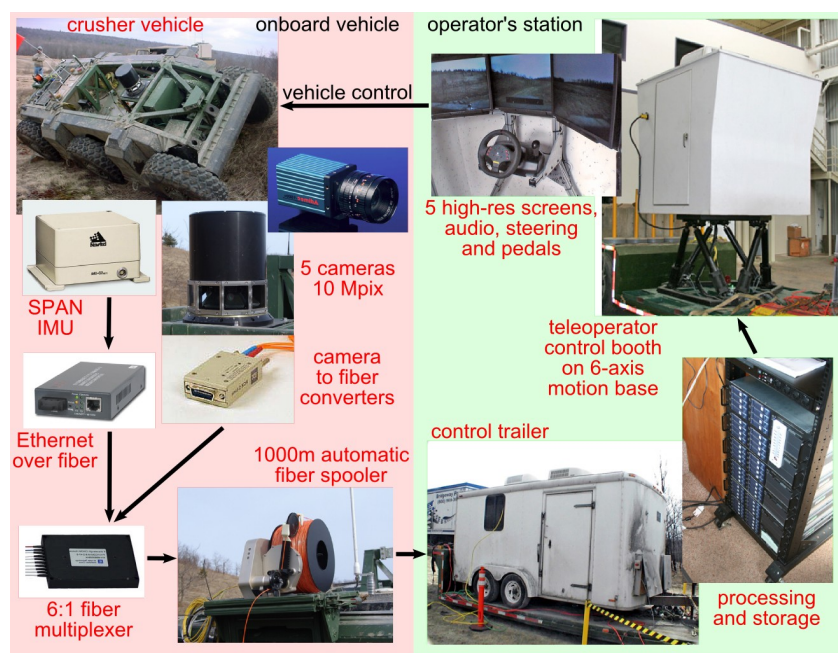


Figure 2: Overall teleoperation system layout

*Control trailer:* The fiber connection from Crusher leads to a control and computing trailer. Five dual-Opteron machines with

high-speed RAID arrays and high-end GPUs process the video feeds and display them in real-time while also saving the unprocessed data to disk. The video is processed using both CPU and GPU programs to perform Bayer filtering, to fine tune the image quality, and to intentionally degrade the image if desired. A high level control program ensures synchronization between the video and audio streams and allows the user to add artificial latency to the system if desired.

*Operator's booth:* The teleoperator rides in a booth mounted on a 6-axis motion base which is capable of greater than 1g of acceleration. The booth includes five video displays which are duplicates of the camera views in the control trailer, stereo speakers, and a standard steering wheel and pedals interface to drive the vehicle.

#### 4. Trial Courses and Methods

We conducted 7 days of testing on the site of an large reclaimed strip mine. This area includes forest, hills, ravines, streams and ponds. During the week of testing we performed 56 trials with a total teleoperated distance of 36 km.

The operator's booth is equipped with a real-time position display on a low-resolution map of the test area. In each trial, the operator was given a goal point on his map and was instructed to reach the goal as quickly as possible. Goals were designed to avoid established paths and roads and to force operators to choose new routes through complex and varied terrain. We used 11 different courses, including simple courses around barrels, courses through thick forest, and courses through rocky ravines. Courses were varied to limit operator familiarization by changing starting positions and goal points, and by placing markers to force operators into new path choices.



Figure 3: Typical test terrain

System parameters, including video resolution, horizontal field of view, frame-rate, and latency, were varied at the start of each run to produce test results based on a wide variety of different levels of perceptual fidelity. Vehicle position, IMU motion, odometry and other data were archived for every trial. The raw video and audio streams were archived to disk for approximately one-third of the trials for later playback and review.

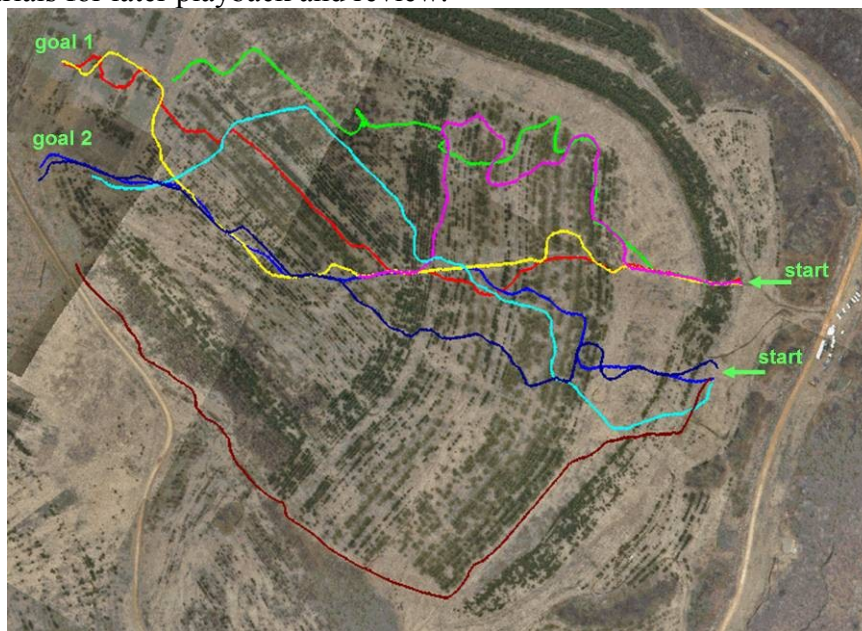


Figure 4: Eight example trial paths plotted from GPS logs

## 5. Results

Due to the limited number of runs which we were able to perform, the complex unstructured terrain, and the large number of variables involved, the statistics we will present must represent a honest best effort rather than a highly rigorous analysis of carefully controlled variables. In our analysis, however, we have relied as much as

Total Latency (ms)	Avg Speed (m/s)	Avg Time Stopped (sec)
360	2.47	64
480	2.20	96
960	1.57	168
1320	1.15	246

Table 1: Latency

possible on quantitative results and in this paper we will only present statistics which are supported by at least 10 trials. Where significant, we will also present subjective impressions reported by the teleoperator and by observers.

*Latency:* Total latency (control lag + perception delay) proved to have the greatest impact on teleoperator performance in our testing. Doubling the total latency from 480ms to 960ms resulted in a 29% reduction in average vehicle speed, and a 75% increase in time the vehicle was stopped by the operator for planning or other reasons. Teleoperator performance declined noticeably, even for experienced operators, with any increase in latency. Operators agreed with NASA Ames results which show [2] that higher latencies force the adoption of a “move and wait” control strategy above a certain latency threshold. While the Ames results put this threshold at 300ms, we felt that drivers of our vehicle were able to adapt to lower levels of latency and were only forced to alter their driving style at levels of 480ms or above. Subjectively, high latency was the factor which teleoperators found most annoying and challenging. Comments such as “latency is horrible” and “latency is killing me” were commonplace.

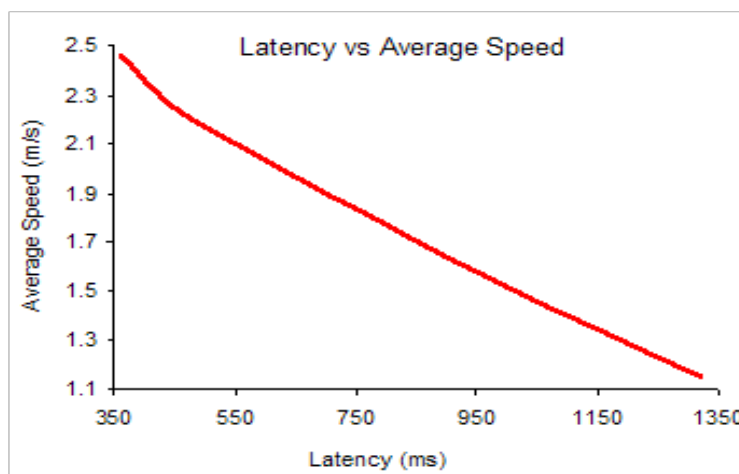


Figure 5: Overall latency vs average vehicle speed  
Graph based on 29 trials

Pixels Per Screen	Avg Speed (m/s)	Avg Time Stopped (sec)
320x240	2.3	138
800x600	2.6	72
1600x1200	3.0	2

Table 2: Resolution

*Resolution:* The impact of display resolution on teleoperator performance was less noticeable than latency, but was still significant. Compared to full

resolution trials (1600x1200 pixels per screen), trials with screen resolutions of 320x240 pixels saw a 23% reduction in average speed and a factor of 69 increase in time stopped. Low resolution video was also found to contribute to poor path planning decisions. In one example, a teleoperator aimed for a “gap” in the forest which turned out to be blocked by trees not visible in the low-resolution view. In another example, a teleoperator reported that he was driving slowly because he could not clearly distinguish obstacles ahead of the vehicle. When the teleoperator was given highest resolution only on the center monitor and low resolution on the remaining four monitors, he reported that he was intentionally yawing the vehicle back and forth to use the higher resolution of the center camera. Subjectively, the higher resolution views seemed to contribute greatly to the sense of realism and presence felt by the teleoperator

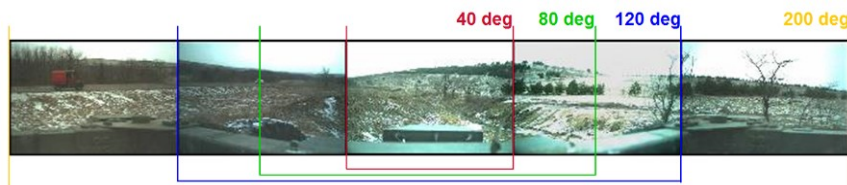


Figure 6: Field of view options

and to increased operator confidence.

*Field of view:* Trials were performed with the horizontal field of view (HFOV) ranging from 40 to 200 degrees and the impact of a limited field of view was found to be significant. Average speed with a 200 degree HFOV was 40% higher than that for trials with a 40 degree HFOV. The time stopped with a 40 degree HFOV was 2 times greater than with a 120 degree HFOV. Paths chosen by teleoperators with narrow fields of view were often poor and several times it was necessary to back the vehicle out of dead ends. Subjectively, field of view had a big impact on the teleoperators’ situational awareness and on their understanding of the vehicle’s position and motion. Several operators expressed a special preference for a wide field of view when the vehicle was moving quickly.

*Frame rate:* Display frame rate was tested at 30, 25, 20, 15 and 10 Hz. Significant effects only appeared at the lower frame rates. For example, average vehicle speed dropped by 37% when frame rate

dropped from 20 to 10 Hz. Teleoperators reported feeling comfortable with 25 Hz but flickering became noticeable and uncomfortable at 20 Hz. Lower frame rates were more challenging and teleoperators reported that it was more difficult to judge vehicle speed when the frame rate was 15 Hz or lower.

*Audio:* Our system included stereo microphones on the Crusher robot and stereo speakers for playback in the teleoperator's booth. Crusher is a hybrid diesel/electric vehicle and is very quiet when running only on batteries, which allowed the teleoperator to clearly hear what was going on around the vehicle. Teleoperators commented that the audio was useful in providing cues such as tree branch and rock strikes, tire slippage, drive motor load and vehicle speed. Subjectively, audio seemed to add considerably to situational awareness and the teleoperators' feeling of telepresence, and teleoperators complained when the audio was turned off.

*Motion:* While the 1g motion base could not faithfully replicate every motion of the vehicle, it could move quickly enough to be uncomfortable when Crusher hit large bumps. Most people who teleoperated the vehicle requested that the gain be turned down on the motion to reduce discomfort over rough terrain. The most experienced teleoperator came to prefer a 30% gain on the motion feedback. Nevertheless, teleoperators reported that the motion was very useful on slopes, over rough terrain, and any time a clear horizon was not visible. Motion sickness was never severe, but was noticeable for some operators.

*Operator's choice trials:* We conducted several trials in which we limited only the teleoperator's total video bandwidth, but allowed him to choose how to spend that bandwidth and to change his allocation at any time during the trial. For example, the teleoperator might choose a single screen at full resolution and full frame-rate, or all five screens at low resolution. This approach proved to be very successful and was well liked by teleoperators. For general driving, teleoperators preferred high frame-rate with medium resolution ahead and low resolution on the sides with a full 200 degree field of view. In complex environments, looking far ahead, and planning routes, teleoperators often traded off field of view or frame rate for higher resolution in the area of interest. This may be a highly

effective method for reducing bandwidth requirements in future teleoperated systems.

## 6. Conclusions

While this paper represents results from only one week of testing, a number of compelling results have emerged. The first observation is that teleoperation with a high-fidelity system is very effective. With the highest quality settings, the feeling of “being in the vehicle” was very strong. Several teleoperators emerged from the control booth believing that they were at the location of the vehicle a kilometer away. Using this system, teleoperators were able to drive effectively, confidently, and safely at the maximum speed of the Crusher vehicle through highly complex and unstructured off-road terrain to the point that safety vehicles, including a HMMWV, often had trouble keeping up.

In many instances the quality of the teleoperation was so good that it surpassed close-proximity visual radio-controlled operation by a skilled operator. We conjecture that the teleoperator benefited from better sightlines provided by the cameras placed directly on the vehicle as well as the physical feedback from the motion base. The subtly, speed, and precision of teleoperation observed in our experiments has convinced the authors that major research advances are required before a fully autonomous vehicle can achieve comparable operation.

This work has also led us to draw some conclusions which may be useful for future teleoperated systems:

- Latency should be minimized and is severe above 400-500ms
- A wide FOV is a great help for situational awareness
- A 20-25 Hz frame rate is acceptable to many operators
- High resolution is essential at long distances in complex terrain
- Audio and motion feedback provide realism and useful cues
- Bandwidth can be reduced by giving the operator choices

## 7. Acknowledgements

Many people went out of their way to help us with this work. In particular, we're especially grateful to: Dave LaRose, Brian Beyer, Jon Chu, Megan Grimm, Mark Waldbaum, Roger Boulet, Mike

Sergi, Don Salvadori, David Rice, Chuck Amurgis, Chris Fromme, and Bill Curtis who all supported this work.

## 8. References

1. Ellis, S.R., Mania, K., Adelstein, B.D., Hill, M. (2004) Generalizeability of Latency Detection in a variety of Virtual Environments. Human Factors and Ergonomics Society 48th Annual meeting, New Orleans, USA
2. Ellis, S.R., Young, M.J. (1999) Adelstein, B.D., Ehrlich, S.M., Discrimination of changes in latency during head movement. Proceedings of the HCI International '99
3. Holloway, R. L. (1997) Registration error analysis for augmented reality. Presence, 6, 4, 413-432
4. Ellis, S.R., Young, M.J., Adelstein, B.D., Ehrlich, S.M. (1999) Discrimination of changes of latency during voluntary hand movement of virtual objects. Proceedings of the Human Factors and Ergonomics Society, September 27-October 2, 1999, Houston, Texas
5. S. Vries and P. Padmos (1997) Steering a Simulated Unmanned Aerial Vehicle Using a Head-Slaved Camera and HMD. Proceedings of SPIE, Vol. 3058, 1997, pp. 24-33
6. E. Kasper, L. Haworth, Z. Szoboszlay, R. King, and Z. Halmos (1997) Effects of in-flight field of view restriction on rotorcraft pilot head movement. SPIE, Vol. 3058, 1997, pp. 34-45
7. McLean, G.F.; Prescott, B.; Podhorodeski, R. (1994) Teleoperated system performance evaluation. Systems, Man and Cybernetics, IEEE Transactions on, Volume: 24, Issue: 5, May 1994 pp. 796-804
8. Volbracht, S.; Shahrabaki, K.; Domik, G.; Fels, G. (1996) Perspective viewing, anaglyph stereo or shutter glass stereo? Visual Languages, 1996. Proceedings., IEEE Symposium on, 3-6 Sept. 1996 pp. 192-193
9. T. Adlin, (1992) Human Factors Issues Associated with Image-Compression for Low Data Rate Remote Driving. Proceedings of SPIE, Vol. 1700, 1992, pp. 489-502
10. Douglas E. McGovern (1990) Experiences and Results in Teleoperation of Land Vehicles. SANDIA REPORT SAND90-0299 UC-515 Printed April 1990