An Examination of the STRIPE Vehicle Teleoperation System

Jennifer S. Kay *
Lockheed Martin
Advanced Technology Laboratories
1 Federal Street
Camden, NJ 08102

Charles E. Thorpe
Carnegie Mellon University
Robotics Institute
5000 Forbes Avenue
Pittsburgh, PA 15213

Abstract

This paper describes a series of quantitative studies of a user interface for robot vehicle teleoperation. Supervised TeleRobotics using Incremental Polyhedral Earth geometry (STRIPE) is a teleoperation system for a robot vehicle that allows a human operator to accurately control the remote vehicle across very low bandwidth communication links, and communication links with large delays.

In STRIPE, a single image from a camera mounted on the vehicle is transmitted to the operator workstation. The operator uses a mouse to pick a series of "waypoints" in the image that define a path that the vehicle should follow. These 2D waypoints are then transmitted back to the vehicle, where they are used to compute the appropriate steering commands while the next image is being transmitted. STRIPE requires no advance knowledge of the terrain to be traversed.

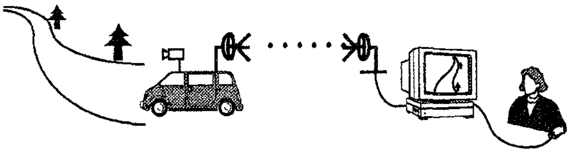
This paper describes a series of tests of the STRIPE system. Individuals with no previous STRIPE experience were given brief instructions in the use of the system, and then asked to use STRIPE to teleoperate the Carnegie Mellon Navlab 2 vehicle along a simple path, through a slalom course, and to find a goal using a very limited on-line map. Conditions tested include different graphical interfaces, bandwidths, lenses, and rates of image compression.

1: Introduction

Supervised TeleRobotics using Incremental Polyhedral Earth reprojection (STRIPE), is a teleoperation system for a robot vehicle [3]. Using STRIPE, a human operator can accurately control a remote vehicle across very low bandwidth communication links, and communication links with large delays.

In STRIPE, the operator is presented with a single image taken from the remote vehicle, and is asked to designate a series of points in that image, known as "waypoints," that indicate where the vehicle should go (Figures 1 and 2).

Once the operator is happy with the points selected, the points are sent to the remote vehicle. The vehicle uses the STRIPE polyhedral-earth reprojection technique [2][3] to follow the designated path. Meanwhile the point picking process is repeated. This is very different from traditional teleoperation methods, in which the operator views a continuous stream of images and directly controls the vehi-



Images are digitized on the vehicle and transmitted to the operator workstation.

Once the image arrives at the operator workstation, the operator uses a mouse to pick waypoints for the vehicle to track.

Figure 1 The basic STRIPE system.

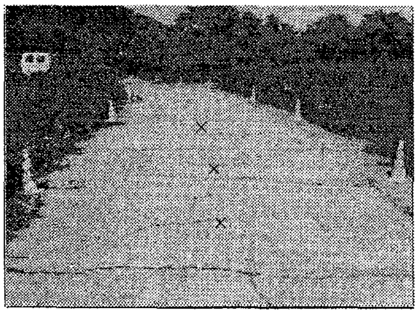


Figure 2 A sample STRIPE image with three points chosen in it.

---

* The work described in this paper was performed entirely while Jennifer Kay was a graduate student at Carnegie Mellon University.

cle's steering using a steering wheel or a joystick. STRIPE can be used in situations where the communication link has a wide enough bandwidth and low enough delay that the traditional methods of teleoperation are also possible. However, STRIPE's strength is its ability to accurately control a remote vehicle when the communication link makes other forms of teleoperation impossible.

This paper describes a series of careful tests designed to provide a better understanding of how the STRIPE system works in practice. The results of the study provide data about the accuracy of the STRIPE system when used by novice operators, and also indicate ways that the STRIPE interface might be changed to improve performance. A more detailed description of these tests and results can be found in [3].

2: The Experiments

2.1: Operators

A total of 14 individuals completed the tests, with ages ranging from 20 to 52. None of the operators had any experience using the STRIPE system before they performed the experiment. Operators received about a half an hour of STRIPE training using prerecorded images, and were then asked to perform three different tasks: teleoperating the Navlab 2 along a simple path, through a slalom course, and to a goal using a very limited on-line map.

We made the decision to use novices for a few reasons. First, we felt that if novices performed satisfactorily, it was likely that expert users of the system could only do better. It was also more practical to use novices. In order to train individuals to become STRIPE experts, they would have needed to have significant practice time using the system, and the setup and run-time costs of this system were very expensive in terms of person-hours. A "safety driver" had to be in the teleoperated vehicle whenever it was under computer control to take over the vehicle's steering in an emergency. Finally, the site was accessible to anyone who wandered by, and so all of the operator control station equipment and obstacles had to be set up in the morning, and brought back to the lab each evening.

2.2: Test Setup

Tests were conducted at a local slag heap. On the majority of the site there was minimal foliage, mostly weeds and the occasional small tree, with some large trees at the edges of the site. Most of the ground was covered with slag, a by-product of the steel manufacturing industry that resembles black gravel.

Tests were performed using the Carnegie Mellon Navlab 2 [4], a U.S. Army HMMWV Ambulance that has been reconfigured to be completely computer controllable (Figure 3). The STRIPE vehicle modules ran on a Sparc 10 mounted on Navlab 2.

The operator controls ran on a Sparc LX class portable workstation, with a display measuring approximately 21.5 cm x 16.5 cm. The display was less than ideal. It was fairly dim, susceptible to significant glare, and displayed only 8 gray-levels in the monochrome images.

The operator control station was located in the back of a cargo van. A card table and two chairs were set up inside of the van for the operator and an observer.

Communication between the operator's workstation and the Navlab 2 vehicle was achieved using a pair of wireless ethernet bridges. On average, 12.9 seconds elapsed between the time the operator clicked the right mouse button to send the waypoints to the vehicle and the time that the next image appeared on the operator's monitor. In addition to the time spent transmitting the data over the wireless ethernet, this value includes overhead due to digitizing, waiting to get current position information from the vehicle controller, JPEG compression, inter-process communication, JPEG decompression, display of graphical pan and tilt values, and display of the image on the operator monitor.

Because of a problem with Navlab 2's automatic speed control at the time of the experiments, the accelerator and brake pedals were manually operated by the safety driver. A speaker mounted over the driver's shoulder would play one sound to indicate that the driver should start to move, and another sound indicating that the brake should be applied. Safety drivers were asked to try to maintain a speed of 0.5 m/s when moving, but there was probably some variation between drivers.

2.3: Course 1: Simple Path Following

The first course consisted of an "obvious" path about 120 meters long. Traffic cones were used to clearly mark the path. The path varied between about 5.5 and 6.5 meters in width. The cones were placed on both sides of the path about every 9-10 meters along the path. Between the last two cones on the path, a yellow tape was laid to designate a stop point. Figure 4 shows a bird's-eye view of the location of the cones for the first path. Figure 5 shows the view

Figure 3 The Carnegie Mellon Navlab 2.

Figure 4 Bird's-eye view of task 1.

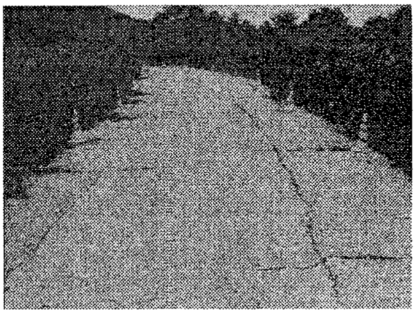


Figure 5 The view from the Navlab 2 camera at the start of course 1.

from the Navlab 2 camera at the starting position of task.

2.4: Course 2: Slalom

The second course was a modified slalom. Figure 6 shows the approximate layout of the cones. Participants were instructed to drive to the left of the single cones and to the right of the pairs of cones (see Figure 7) and were told that the cones are spaced unevenly. The course was designed to become progressively more difficult. It began with an gap measuring approximately 14 meters between the first two obstacles (a single cone and a pair of cones), and then approximately 14, 12, 12, 10, 10, and 9 meters between subsequent obstacles. It was not expected that any of the operators would be able to complete the entire course.

The cones were offset to the left and the right to make the task possible: camera pan was limited to approximately 36 degrees from center and this meant that cones were often out of view, even when the camera was panned to its maximum position. Figure 8 shows a view of the course from the Navlab II vehicle.

Figure 6 Bird's-eye view of task 2.

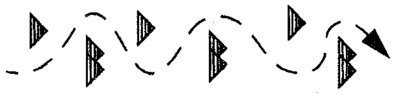


Figure 7 Operators were instructed to drive to the right of the single cones and to the left of the double cones.

Figure 8 The view from the Navlab 2 camera at the start of the second course.

2.5: Course 3: Map Following With Obstacles

The final course made use of a simple map interface shown in Figure 9. The map shows the vehicle, and a vector pointing towards the goal (Figure 9a). The horizontal axis of the map is parallel to the real world x-axis, and the vertical axis of the map is parallel to the real world y-axis. Every time a STRIPE image is delivered to the operator, the map is updated: the old vehicle location is shown in gray and the new location is drawn in black (Figure 9b). If the vehicle runs off the edge of the map, the visible area covered by the map is shifted so that the vehicle is drawn in the center. When the goal is located on the visible area of the map, it is displayed as a star (Figure 9c).

Operators were instructed to drive to the goal, which was designated by a group of three cones, and was out of the initial view of the camera. The direct route to the goal, about 70 meters from the start point, was blocked by two pairs of cones with triangular flags strung between them (Figure 10). The initial image from the vehicle at the start of task 3 was uninformative (Figure 11). The purpose of the task was to see how well participants made use of the map interface to find the goal.

2.6: STRIPE Interfaces

All of the fourteen operators used the STRIPE system

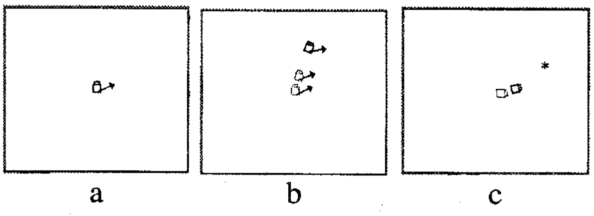


Figure 9 Different views of the map interface as the vehicle heads towards a goal.

Figure 10 Bird's-eye view of task 3.

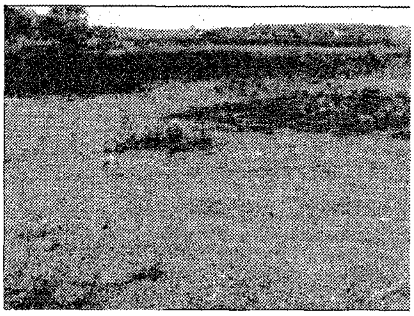


Figure 11 The view from the Navlab 2 camera at the start of task 3. No cones are visible.

individually, but the interfaces they used were divided into seven categories.

The *standard* interface used the default STRIPE interface: an 8mm lens, and experienced a delay of about 12.9 seconds between sending points and the new image being displayed.

The *reduced bandwidth* interface had a narrower bandwidth than the standard interface, and had a delay of about 18.2 seconds between sending points and the new image being displayed.

Glumm et. al. [1] demonstrated that for certain high-bandwidth teleoperation tasks, different fields of view were preferred. We were interested in finding out whether these results held for low-bandwidth teleoperation tasks. The *wide field of view* interface used a 3.6 mm lens, which introduced some significant radial distortion. The *narrow field of view* interface used an approximately 12 mm lens. Figure 12 shows images digitized using the two lenses.

The *reduced resolution* interface had a higher rate of image compression, with image that took up about half the space of the standard images. Despite the significant reduction in data, the low quality of the operator workstation monitor meant that the resulting images were not that much worse than the standard images. The images only appeared at the operator workstation about a half a second faster than their larger counterparts due to overhead.

The *dashboard* interface used a graphical pan/tilt interface that attempted to display the pan as a view through a virtual windshield.

The *no graphical pan/tilt* interface only provided a numeric value for pan and tilt.

a b

Figure 12 The view through the narrow (a) and wide (b) field of view lenses at the start of the first test course.

3: Course Results

3.1: Course 1: Simple Path Following

All but one of the operators completed the first course successfully, without hitting any of the cones, on the first attempt. Most of the operators reported that the first task was fairly simple.

It is important to note that this was the first time each operator used STRIPE to control a remote vehicle. The training they received before attempting this task was done using prerecorded images, presented in a fixed sequence. The position of a user's waypoints in the current image had no affect on the next image that was displayed. Nevertheless, 13 out of 14 operators successfully navigated the remote vehicle through the first test course the first time they used the real STRIPE system, despite the poor reproduction of images on the operator workstation monitor.

3.2: Course 2: Slalom

The slalom course was designed to start out relatively simple, and become increasingly complex. It was not expected that any of the operators would be able to complete the course. The slalom was also designed to force operators to pan the camera on the Navlab 2 vehicle, and it succeeded. In the first course less than half of the participants used the pan, and most only used it twice. In the second course only three participants did *not* use the pan.

All of the operators successfully maneuvered around the first two, relatively mild obstacles on the slalom course. More than three quarters of the operators successfully completed the increasingly complex first half of the course, and one operator succeeded in completing the entire, extremely difficult, course. Most of the operators found the slalom to be the most difficult of the three tasks.

3.3: Course 3: Map Following With Obstacles

Over two thirds of the operators successfully completed the map following task. It took some of the operators a few STRIPE cycles to get used to the map interface. The map that they were shown during training started with the vehicle pointing straight upward, as in Figure 9a. During the actual task, the orientation of the vehicle on the map was determined by the vehicle yaw. The inertial sensor usually initialized a yaw of 0 to be North, and so the initial yaw of the vehicle was usually about 240 degrees. Thus the vehicle on the map was pointing downward and slightly to the right. This confused some of the operators, who initially believed that they had to turn the vehicle around so that it was pointing upwards on the map before they could reach the goal. After seeing an updated map with the new position of the vehicle relative to the old one, most operators recognized their mistake.

3.4: Surprises

3.4.1: Image Digitizing and a Moving Vehicle. By far, the most common misunderstanding among operators when they began to use the system had to do with when new images are digitized.

STRIPE was designed to continue moving as long as it has waypoints ahead of the vehicle. When a set of points is sent from the operator to the vehicle, the vehicle takes a new image, and sends that back to the operator. Meanwhile, the vehicle continues to follow the path.

The important point to note is that the image the operator receives is often "old" by the time they receive it, i.e. the vehicle is no longer in that position. This is not a problem for the STRIPE system, points that are behind the vehicle are ignored.

This was explained to operators during their training. Nevertheless, a large number of operators expected that an image would not be taken until the vehicle came to a stop, and expressed surprise when an image that they received did not meet their expectations.

As they continued to use the system, several operators began to understand what was happening. By the end of the last task, half of the users had started requesting new images from the vehicle, without sending any points to the vehicle. Since they knew that the image that they had just received was probably not the "final" image, they would ask for a new image frequently, to get an update of where they were in the task. Of course sending a new image across a limited bandwidth link is very time consuming and might have slowed some of the operators down.

STRIPE was carefully designed to send images while the vehicle was still in motion. It enables operators to pick points while the vehicle is in motion, which could enable the vehicle to move a longer distance in the same period of time. This decision will have to be reassessed in future implementations of STRIPE.

3.4.2: Image Digitizing and a Stopped Vehicle. Before the user studies began, it was understood that the first two images in a STRIPE run always looked almost identical. This happens because the second image is generally digitized before the vehicle has moved very far using waypoints designated in the first image.

The vehicle is initially stationary and the first image is sent to the Operator Workstation. The operator's 2D waypoints are sent to the STRIPE Main module, and the Image Capture module is informed that it is time to digitize a new image. If the image capture module were to digitize an image immediately, the vehicle would not have moved at all, and the first two images would look identical. To prevent this from happening, STRIPE was adjusted to pause for up to two seconds if the vehicle and the camera position had not moved since the previous image was digitized. Two seconds is not really much time when the speed is being controlled by a safety driver who has a certain delay between hearing the "start vehicle" sound, and actually starting to move.

This "duplicate image phenomenon" occurs whenever the vehicle runs out of waypoints from image n, and comes to a stop at a point beyond the last point on the next set of waypoints from image n+1 that have not yet been received. Thus the STRIPE tests were designed with a very slow moving vehicle which, it was thought, would eliminate this problem.

What was not anticipated was how hesitant certain operators would be about picking points. Operators were instructed to be careful about the points chosen, and many were extremely careful, picking very short paths. This, when combined with the limited bandwidth link, meant that several operators received duplicate images in the middle of their runs.

3.4.3: Images From Unexpected Orientations. Some operators expressed the feeling that the image that they were looking at was taken from a completely different angle than they expected, and had nothing to do with the path that they had picked. Initially it seemed as though perhaps the operators were correct, and that the vehicle had somehow driven far off course. But this was not the case, the vehicle was still on track. What had happened to produce this?

Imagine that the "duplicate image phenomenon" described in Section 3.4.2 has occurred, and so an operator picks two sets of points in what is essentially the same image. Figure 13a shows an overhead view of the locations on the ground of those points. The first set of points, the gray ones, are chosen and sent to the vehicle and the vehicle begins to drive. When the next, identical image arrives at the operator workstation, the user picks what appears to be approximately the same path, the clear circles. Suppose that when the new set of points arrive, the vehicle is positioned just at the second gray circle. When the clear points arrive, the vehicle is suddenly informed that it is several centimeters to the left of the path that it should be on. It makes a sharp turn to the right to get closer to the right path, and then almost immediately a correcting sharp turn to the left, see Figure 13b. The position of the rear axle of the vehicle, the location of the origin of the vehicle, does not move too drastically. But any images digitized just as the vehicle is pulling over to the right will look as though the vehicle is headed well off the designated path. Notice that the first two images do not have to be identical for this to happen. The problem can occur whenever the new path that the operator designates is off to one side or the other, even slightly, of the previous path.

3.5: STRIPE Interfaces

A large amount of data was collected during the STRIPE experiments. This data included information about the distance travelled, time elapsed, where in the images operators picked points, the number of images transmitted and their orientation, how frequently the vehicle came to a stop, and more. Much of this data is analyzed in [3]. Because there were only two operators with valid data in each category of interface, it is not possible to show that any of the interfaces was statistically significantly better than any of the others. We can, however, examine in more general terms how operators in each of the seven categories fared.

The reduced bandwidth condition has the potential to help operators see the images they expect to see some of the time. The fact that sending images takes more time means that the duplicate image phenomenon is more likely to occur, because the vehicle has more time to run out of points in a path. If the operator recognizes that the new image is a duplicate of the previous one, and picks about the same points in the duplicate image, then there is a good chance that the next image that the operator will see will be at the end of that path.

One big problem with using a narrow field of view lens is that any rotations of the vehicle are emphasized, because the operator can not see expected landmarks. In contrast, the wide field of view lens does allow operators to see more of the context of what was going on around them, but objects appear farther and smaller in the images.

Operators seemed to understand the standard pan/tilt interface better than the dashboard and numeric ones.

Because of the poor quality of the operator's monitor, the reduced quality of the images that were more heavily compressed was not as significant as it might have been. This interface was probably not very different from the standard interface.

4: Conclusions

The STRIPE experiments demonstrated that novice operators can use the system effectively with only minimal training. The experiments also provided a wealth of information about how users expect the system to work. This paper provides a brief overview of a few of the most interesting of these results, a complete description can be found in [3].

The differences between operators' expectations about the system and the actual performance of the system highlight those areas that have potential for improvement. This information will be used in the design of future implementations of STRIPE.

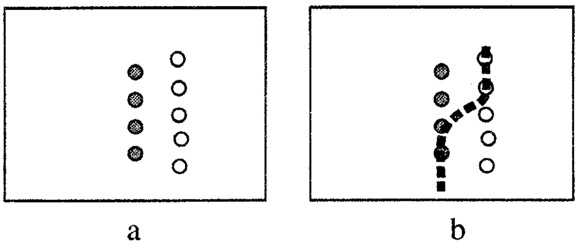


Figure 13 An overhead view of the real world locations of two sets of points chosen in two identical images (a), and the vehicle path through those points (b).

The STRIPE system is a unique combination of computer control and intelligent user interaction. The computer determines the 3D world path designated by the 2D waypoints and then accurately controls the vehicle over rugged terrain. The human must select points in the 2D image that accurately designate where the vehicle should go. The STRIPE experiments have demonstrated to us the value of careful user testing for robotic systems that require human input.

5: References

[1] Glumm, Monica M., Kilduff, Patricia W., and Masley, Amy S., "A Study on the Effects of Lens Focal Length on Remote Driver Performance," *ARL-TR-25*, Army Research Laboratory, November 1992.

[2] Kay, Jennifer S., and Thorpe, Charles E., STRIPE: Supervised TeleRobotics Using Incremental Polygonal Earth Geometry, Proceedings of the International Conference on Intelligent Systems 3, Pittsburgh, PA, February 1993.

[3] Kay, Jennifer S., STRIPE: Remote Driving Using Limited Image Data, Ph.D. Dissertation, Computer Science Technical Report 97-100, Carnegie Mellon University, Pittsburgh, PA 15213, January 1997.

[4] Thorpe, Charles E., ed., Vision and Navigation, *The Carnegie Mellon Navlab*, Kluwer Academic Publishers, Norwell, MA, 1990.

6: Acknowledgments

Jennifer Kay was funded by a NASA Graduate Student Researchers Program Fellowship grant number NGT-51292. This research was also partly sponsored by DARPA, under contracts "Technology Enhancements for Unmanned Ground Vehicles," contract number DAAE07-96-C-X075 and "Unmanned Ground Vehicle System" contract number DAAE07-90-C-R059 monitored by TACOM; and "Perception for Outdoor Navigation," contract number DACA76-89-C-0014 monitored by the US Army Topographic Center. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the funding agencies.