4

Remote Control

When a direct connection exists between operator and remote sites, then bandwidth is virtually unlimited (if you can run one wire, you can easily run another), and there is no perceptible communications delay. Under these conditions there is no cost to sending information between sites. Its worthwhile transmitting every video frame and every bit of force feedback just in case it might be helpful.

When the direct wire is replaced with a more indirect connection, such as the Internet, then the nature of communication changes. Now there is a limited bandwidth and communications are delayed. In addition, since it is a shared resource, there is an advantage in not sending information, even when it may be useful at the other site.

This chapter introduces the reader to the problems of constrained communications. By way of explanation, we'll pick some examples of long-range teleoperation from the realm of the Internet and, in particular, the World Wide Web. Implementations in this domain must cope with delays on the order of a few hundred milliseconds and bandwidths on the order of a few hundred kilobits per second. These systems are relatively new and rather simple in comparison with traditional teleoperation; however, they have the advantage of being accessible to millions of users. As in life, popularity is often more important than functionality.

We'll begin with the transmission of images from a remote site and then move on to consider the control of remote machines and sophisticated robot manipulators.

4.1 Control of remote cameras

The control of remote camera systems via the constrained communications of the World Wide Web is the most popular remote control technology available via the Internet.

Early implementations allowed remote users to watch a coffee pot at the University of Cambridge[1] and a fish tank at Netscape Communications Corporation.[2] One excellent recent example is the VolcanoCam—an Internet camera that allowed users throughout the world to view live images of Mount Ruapehu, an active volcano in the central North Island of New Zealand.[3] Figure 4.1 shows the operator interface. Users request a new image by simply clicking on the old picture.

The VolcanoCam takes a new image every minute and uploads it via a cellular modem to the Web server. What happens after that is constrained by the limited bandwidth of the Internet.

Dealing with bandwidth constraints requires three steps:

- Reducing the information content.
- Encoding it efficiently.
- Sending it infrequently.

The VolcanoCam's designers reduced the information content by choosing a slow frame rate and a small frame size. They encoded each frame efficiently (and further reduced the information content) by using a lossy JPEG image compression algorithm [62]. Finally, they reduced the amount of information transmitted by sending images only on demand.

The use of compression allows us to trade off between the computational burden at each site and the bandwidth requirements between sites.

---

[1] http://www.cl.cam.ac.uk/coffee/coffee.html
[2] http://www.netscape.com/fishcam/fishcam_about.html
[3] http://www.cybercorp.co.nz/ruapehu

"... And then I started to think. I could now go home, open my laptop, dial in over two loosely bound pre-historic copper cables, and get live pictures in my living room of a mountain four hours' drive away. "

—*Julian Meadow*

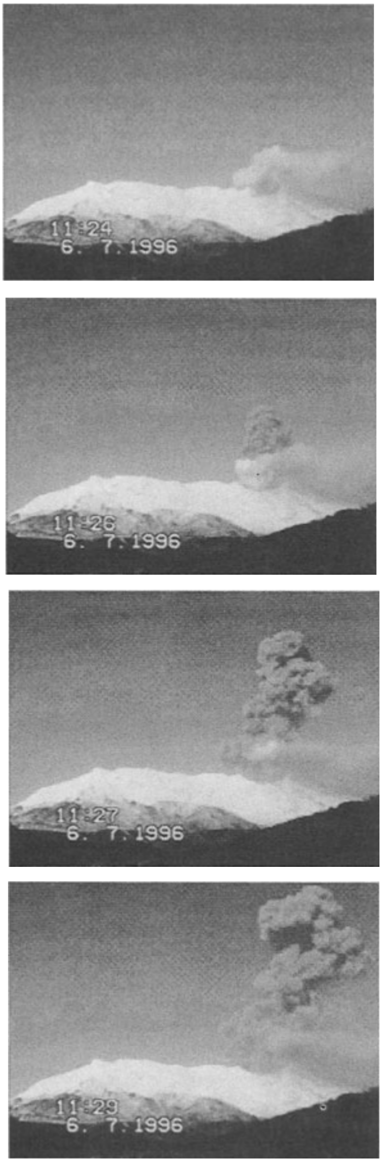


FIGURE 4.1: The VolcanoCam is one of the first, and certainly most unique, examples of transmitting live imagery from a remote site via the Internet using a WWW user interface. The interface is shown on the right. Shown at left are the images that users saw on July 6th, 1996, while Mount Ruapehu was erupting. Thanks to Julian Meadow and CyberCorp (NZ) Ltd, for allowing us to reprint these images.

The communications delays imposed by Internet communication are just as much a problem as the limited bandwidth; unfortunately, they are much harder to overcome. One technically possible design is to send images continuously to anyone who happened to have the VolcanoCam web page open. In this case, communication is unidirectional and, aside from the delay for the first image, communications delays have no discernible effect on the user's experience.

However, bandwidth on the Internet is currently in high demand and so, while this is technically possible, it's socially inadvisable. Instead the system only sends a new image when the user explicitly requests one.

In the case of the VolcanoCam, all tradeoffs were made at the design stage and hard-coded into the system. In later chapters, we'll look at some mechanisms for dynamically trading off limited resources.

Of course, the viewing of live images from a remote site is not, in itself, remote control—one cannot influence the outcome of a live television broadcast merely by sitting in front of the screen—however, visual imagery is a powerful and very natural means to transfer information on the state of the remote site to the operator, and it will form a part of most remote control implementations.

4.2 Controlling a remote machine

Perhaps the most interesting examples of remote control via the Internet are those that allow users to actively interact with a remote environment. Here the effects of delays become more pronounced and the solutions more demanding.

In this section, we'll look at three different examples of remote control using a World Wide Web interface. The first is a robot archaeologist, the second a toy railroad, and the third a six-degree-of-freedom industrial manipulator.

4.2.1 Robot archaeologist

One of the earliest implementations was the Mercury project at the University of Southern California.[4] This allowed operators to explore a toy

[4] http://www.usc.edu/dept/raiders/

archaeological site by controlling a remote robot. The user could direct the robot through a natural Cartesian interface and command it to blow puffs of air, disturbing the sand to reveal artifacts hidden at the remote site.

This system was online from August 1994 through March 1995. In that time, over 2.5 million hits to the site were registered.

The operator interface for the Mercury project is simple and elegant (see Figure 4.2). Users interact with two pictures. The first shows an image from a camera on the remote manipulator. Here the operator indicates the point in the current image that they would like to be in the center of the camera view after the motion. For larger moves, where the desired end-point may not yet be visible on-screen, there is a second image showing a top-down view of the entire robot workspace. Once again, the operator merely clicks on the point to which they would like the remote manipulator to move. In addition to the clickable maps, the operator may raise or lower the remote arm and may command it to blow a puff of compressed air. The interface works well because most motions are confined to the horizontal plane. The system is reliable because its interaction with the environment is limited to puffs of air.

Once again, constrained bandwidth communication is accommodated by sending small compressed images. The elegant interface allows relatively high-level motion commands to be sent from the operator to the remote site. The limited interaction with the remote environment makes it unlikely that unexpected events will occur—there's only so much damage you can do with a brief puff of air.

4.2.2 Toy railroad

Another example is the Interactive Model Railroad from the University of Ulm in Germany.[5] This allows anyone with a World Wide Web browser to control a model railroad and observe the result (see Figure 4.3).

In this case, the limitations of delayed, low-bandwidth communications are handled by sending relatively high-level commands from the operator to the remote site and by sending small, compressed images back. Unlike a home train set, it's not possible to vary the locomotive's speed or suddenly

[5] http://rr-vs.informatik.uni-ulm.de/rr/

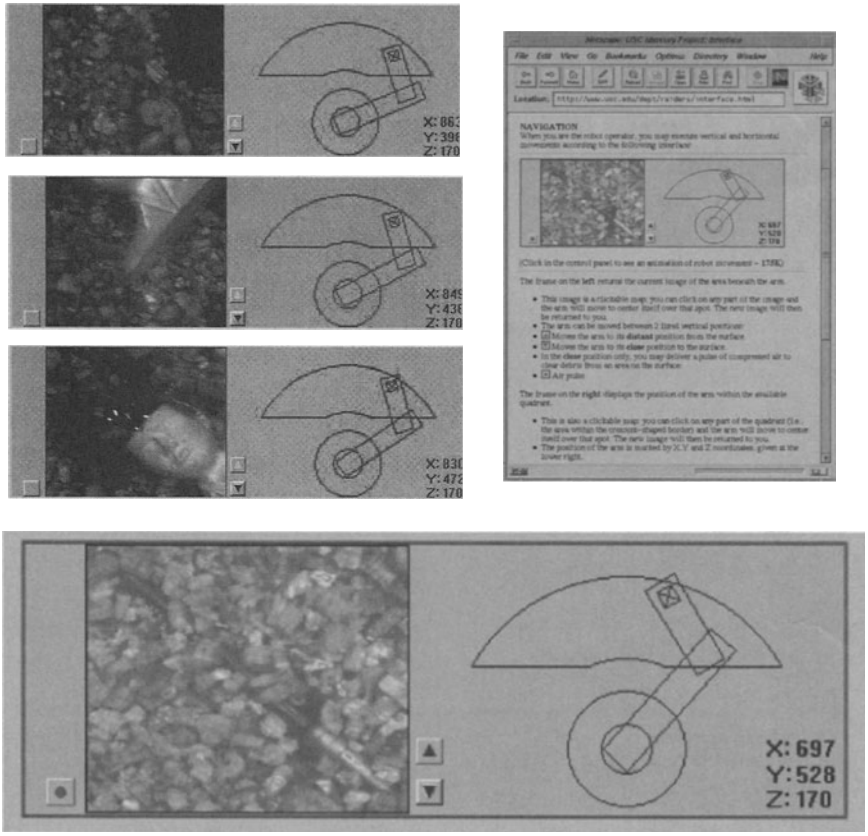


FIGURE 4.2: The Mercury Project at the University of Southern California allowed remote visitors to explore an archaeological site. The three images at top left show an example motion; the full interface is shown at right with an expanded view below. Note the two clickable images that form the primary operator interface. The left image shows a view from the remote arm-mounted camera, while the right image shows a stylized view of the robot's workspace. Thanks to Ken Goldberg for allowing us to reprint these images of his system.

FIGURE 4.3: The Interactive Model Railroad goes one step further than the VolcanoCam—it not only shows users images of a remote site, but also allows them to interact with it—controlling the operation of a model railroad. Thanks to Heiner Wolf and Konrad Froitzheim of the University of Ulm, Germany, for allowing us to reprint these images of their system.

change direction. Instead, one sends commands of the form "make train 2 go to platform 3."

Here the designers have raised the level of communication between the operator and the remote site. Instead of sending dozens of messages a second, the system instead is transferring only one command every few seconds. This idea—of overcoming bandwidth constraints by communicating at a more abstract level—is fundamental to remote control via constrained communications.

Of course, since the trains are real objects, they are unfortunately subject to physical constraints. They may collide or derail and, in such cases, while the operator may be able to see the problem, he or she is unable to do anything about it.

The difficult part of remote operation is not the transferal of commands to the remote site, nor is it the interpretation of those commands as physical actions. Instead it is the development of mechanisms to detect and overcome those problems that will inevitably occur. The need to work despite failures will be revisited many times in this text.

4.2.3 An industrial manipulator

Researchers at the University of Western Australia have made a five-axis industrial robot available for use on the Internet via a World Wide Web interface.[6] Here the remote site is considerably more complicated, the robot has six degrees of freedom (five motions plus the gripper), it is observed by four cameras, and it has a table full of toy building blocks within its workspace.

The operator interface (see Figure 4.4) shows views from each of the remote cameras and accepts commands in the form of Cartesian end-effector coordinates. This is more powerful, but somewhat less elegant, than the previous examples.

Once again, small compressed image fragments are sent from the remote site to the operator. However, in this case the operator has a number of manual controls in order to trade off between the amount of visual information and the rate at which it is updated. He or she may disable cameras, vary the size of each image, and even vary the number of discrete grey levels used to encode each image.

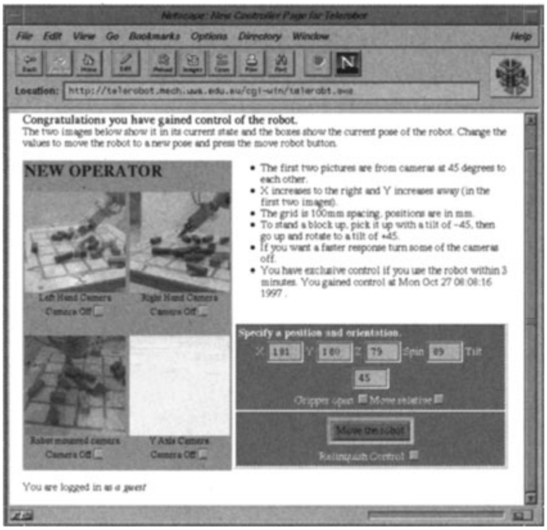
This system provides the operator with manual control of every detail, from the position of the end-effector to the number of greyscales used to transmit each image. This gives him or her considerable power and flexibility. However, that power is not without cost. In some sense, every additional control imposes a tax on the operator.

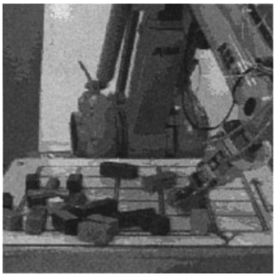
One solution is to follow the lead of conventional teleoperation by adopting predictive displays and automated camera control. Another is to raise the level of communication between operator and remote sites, providing the remote site with enough information to enable it to make some decisions autonomously without needing to bother the operator. A system that does this is the subject of the next chapter.

4.3 Summary

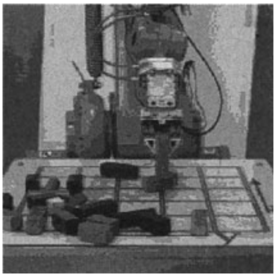
The limitations imposed on the VolcanoCam by Internet communication—low bandwidth and delayed communications—will occur repeatedly through-

[6] http://telerobot.mech.uwa.edu.au

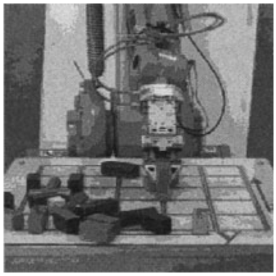




Begin with gripper open, $x = 150$, $y = 150$, $z = 7$, $spin = -45$, and $tilt = 45$.



Close the gripper, move up by setting $z = 100$, then reorient gripper with $x = 200$, $y = 200$, $spin = -89$, and $tilt = -45$.



Now lower the block by setting $z = 37$.

FIGURE 4.4: The University of Western Australia has one of the most sophisticated robots on the Internet. Shown here is the user interface for commanding the robot and an example sequence of commands for repositioning a block. For more information see http://telerobot.mech.uwa.edu.au. Thanks to Ken Taylor for allowing us to reprint these images of their system.

out this book. In this case, they are handled by transmitting images only on demand, by choosing a small image size, and by using JPEG compression.

Existing systems for teleoperation via the Internet range from the simple and elegant (the Mercury project), to complex and powerful (the UWA robot). Ideally we'd like a system that provides the simplicity of the former while still providing the power of the latter. The key to achieving this is to build on existing research in conventional teleoperation and to raise the level of communication between sites.

It's certain that teleoperation via the Internet will improve significantly over the next few years. Some advances that can be expected include overlaying visual clues, showing predicted motions, and automating the control of remote cameras.