

Sensory Feedback and Remote Control of Machines in Mining and Extraterrestrial Environments

By Barry Richardson, Dianne Wuillemain and Mark Symmons
Bionics and Cognitive Science Centre
Monash University, Churchill

ABSTRACT

Remote control can offer some advantages over full automation, however much information is lost and performance suffers if the controller's feedback is restricted to a standard TV display. To restore more normal performance, an operator can be given a stereoscopic view of the work area along with haptic (touch) and auditory feedback. It is also possible to provide enhanced feedback to improve performance beyond that possible in the "real" situation. Lessons learned with remotely controlled mining machines should be applicable to remotely controlled vehicles in different unpredictable environments, such as other worlds.

AUTOMATION AND REMOTE CONTROL

Fully automated machinery is still best suited to relatively repetitive tasks requiring little or no adaptation to changes in the environment. With neural networks and other advances in machine intelligence this is changing, but for now, remote control offers two distinct advantages. The first is that it allows us to exploit the inbuilt flexibility and adaptability of the human operator so that unexpected demands can be accommodated. The second is that experience gained with remotely controlled machinery can teach us relatively inexpensive lessons about the features a fully automated machine must eventually have.

PROBLEMS WITH REMOTE CONTROL

Miners and extraterrestrial explorers share an interest in rocks and minerals. For ore extraction, large rocks containing the ore are manoeuvred by a rockbreaker to allow a pointed spur to be positioned strategically and driven in with sufficient force to split the rock. Pieces must be small enough to pass through sizing grids before being moved, often from an underground tunnel, to the surface. One mining manager describes the smooth operation of a rockbreaker as "poetry in motion", but removed from such machines and given remote controls, the efficiency of the operator can be markedly reduced (Rastogi, Milgram, & Drascic, 1996). This is often due to less than optimal vision, lack of haptic feedback, and sometimes reduced auditory information (Roberts *et al*, 2002; Matti, 2002). Stereophonic sounds that faithfully reproduce the audible work environment are not difficult to arrange, but the provision of vision in depth, and haptic information (kinesthesia and touch) are more challenging goals.

Analogous problems exist for remotely controlled vehicles (ROVs) in extraterrestrial settings. They must be able to manipulate and collect rock samples and be flexible enough to cope with additional and unexpected challenges related to the novelty of the environment.

IMPROVING DEPTH PERCEPTION

Although depth perception relies on many sources of information, one of the most powerful is retinal disparity that results from having two eyes and therefore seeing slightly different aspects of an object, or objects (Palmer, 2003). The greater the disparity, the nearer the object, although we are not conscious of using this information in judging depth. Three-dimensional displays in movies deliver a different image to each eye of an observer. This difference mimics the

retinal disparity present in normal binocular vision and has been achieved in the past with glasses worn by the observer (Goldstein, 2002). Each lens of the spectacles has a filter of some kind (eg polariser, colour, or shutter) to produce the disparity, and the brain fuses the images into one to yield a percept in which extent of disparity is inversely proportional to perceived depth (or distance). Systems using such spectacles and displays have been developed to improve motor performance in dynamic telepresence environments (Bradshaw *et al*, 2002).

It is now possible to create a three-dimensional autostereoscopic display on a flat-screened monitor without the need for such glasses (Schmit, 2002). Essentially, the right and left eyes receive disparate views because the source image on the screen is covered with an invisible vertical grid that separates the two views. Alternate columns of pixels in the monitor project to the right or left eye only. The different views are recorded by two or more cameras, the outputs of which are combined into "double image" displays by purpose-built software. Recent versions have overcome the need to view the monitor from straight ahead only (the single "sweet" spot) and can now be viewed from several directions (Schmit, 2000).

If cameras are mounted at key locations on an a rockbreaker, as they are on the vehicles exploring Mars at the time of this writing, a remote operator of that machine would be able to see the target material and relevant machine parts in a three-dimensional display. This would be expected to improve vision significantly and also to overcome the negative reactions to 3D spectacles that may interfere with normal vision of objects surrounding the monitor, including hand controls.

Though likely to vary according to conditions, the number and placement of cameras needed to display a target in sufficient depth can be determined empirically. A single camera with two lenses is likely to be the best option, and such a system has been used effectively for underwater exploration (Woods, 2003). In some well-lit circumstances without "clutter", stereoscopic vision may not be necessary, but when many rocks of different sizes partially obscure one another, it is likely to be crucial. Another empirical question is whether machine-mounted cameras will vibrate intolerably, or with too much "jiggle" to be electronically suppressed. Wall mountings or portable tripods may be preferable in some terrestrial settings.

Three-dimensional vision may well be *more* important in extraterrestrial exploration and mining environments than it is in normal ones, because monocular cues will be reduced. Powerful monocular depth cues include previous experience of the size of objects, both alone and in relation to others, and partial occlusion of one such object by another. However, many environments in which machines need to be remotely controlled have impoverished monocular cues. For example, there is no "standard" size of a rock to be learned, so the retinal size of an object and partial occlusions of such objects are not reliable depth cues in unfamiliar settings such as underground tunnels. In contrast, retinal disparity cues do not require such experience, and are probably innate (Palmer, 2003).

Sensory Feedback and Remote Control of Machines in Mining and Extraterrestrial Environments

IMPROVING HAPTIC PERCEPTION

Remote users of rockbreakers complain that they no longer "feel" what is going on. This problem could be addressed with the use of force feedback devices (or adaptations of them), such as the *Phantom*, produced by *Sensable Systems* in the USA. The user of a *Phantom* holds a probe that is connected to linked shafts through several joints (with a total of six degrees of freedom) that sense the position of the probe in relation to a virtual object, the surfaces of which are defined by spatial coordinates in a computer. When the probe "reaches" the virtual surface, resistance is felt in a manner analogous to a blind person sensing resistance at the end of a cane when probing an object with it. In the *Phantom*, the resistance is provided by motors located at the links within the multijointed device and these motors can also "drive" the probe, and therefore the hand holding it, as would happen if the object being explored is elastic, or in motion so as to push the probe. Virtual objects of great variety, complexity, and texture can be computer-generated, explored, and moved in haptic space within limits set only by the software that defines the object's characteristics, and the capacity of the human kinesthetic (limb and joint position) system to detect and interpret the inputs. Recently, two *Phantom*s have been used by people in different countries, sharing the same virtual space, to "shake hands" over the internet.

This technology can be adapted for use on a remote machine that has force feedback devices fitted to sense movement and vibration of controls, or other parts, on it. These movements and vibrations can be transmitted to controls on the surface so that the operator can feel the same haptic sensations that are normally available. The operator's manipulation of the controls would not only cause corresponding movement of the machine, they would cause sensors on the machine to haptically feed back the consequences of the operator's commands. This haptic feedback would combine with the visual channel to offer improved perception and performance.

Like its mining counterpart, an ROV may have controls that allow it to be driven by an astronaut wearing a space suit, or remotely controlled (for greater durations and distances) from a base. Alternatively, the machine may not be "drivable" at all. The controls could be present at the remote site only, with

commands and feedback signals received and transmitted by specialised systems on the machine. Either way, it will be possible to enhance sensory feedback by, for example, including variable mechanical resistance on a lever that controls hydraulic arms. If this pressure were proportional to, say, the weight of the load, as it would be with purely mechanical linkage, a valuable loop is created. Such feedback is often lost with hydraulically powered machinery. In general, the contribution of haptics information to these and "augmented reality" environments has been neglected (Rauterberg, 1999).

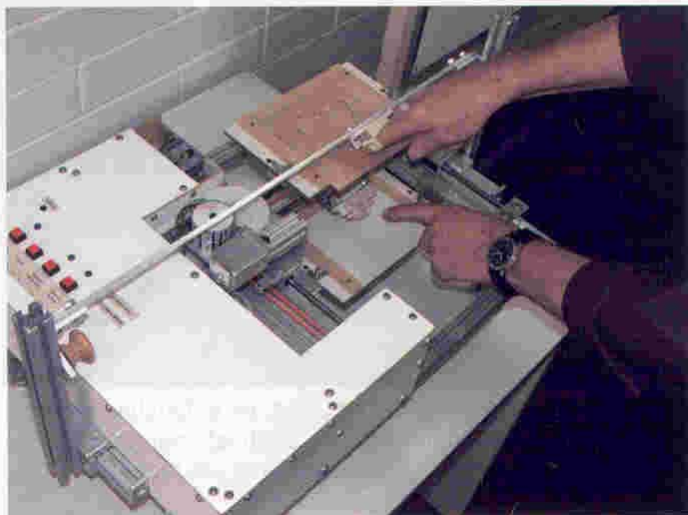
SENSORY DOMINANCE AND CAPTURE

The shape of an object can be perceived visually or haptically, and with normal objects vision usually dominates or "captures" information from touch or kinesthesia that is discrepant with it (Shiffman, 1996). Touch rarely captures vision – although the sharpness of a knife is best tested on the skin, counterfeit money is often discovered because it feels wrong, and the haptic sense can dominate when visual acuity is reduced (Heller, 1983). We are currently investigating the extent to which discrepancies in the visually and haptically perceived shape of *virtual* objects lead to sensory capture, and to do this we have had to design programs to interface the haptic and visual stimuli. By manipulating how well the visual and haptic images agree, we can measure tolerance for a degraded image in one modality given higher fidelity in the other. This kind of knowledge will prove useful in assessing the value of haptic feedback when vision is poor such as during dusty conditions or low light.

THEORETICAL ISSUES

An implicit argument in the above is that multisensory displays are better than unimodal displays (Sarter, 1998). However, to take advantage of the benefits of multimodal signals, it is not enough to simply "throw in" other stimuli or dimensions. The relationship between the inputs is important if facilitation is to be achieved, and interference avoided. Immersive environments that are rich in multimodal inputs can be overwhelming and confusing, rather than beneficial. In general, congruent or redundant inputs complement each other and may substitute for one another to improve perception. However, optimal human-machine interfaces remain elusive and it is anticipated that in studying the benefits of multisensory inputs to the operator of a remote mining machine, much will be learned about optimising teleremote displays in other settings such as surgery, bomb disposal, and training simulators.

We plan to look for advantages in including as much sensory information as possible in order to approximate normal conditions in which attention is directed to *selected* features of complex arrays containing, as they typically do, many redundancies and potential distractors. The brain does the selection and is suited to this task, which it performs at an unconscious level. This approach contrasts with attempts to *simplify* stimulus arrays with the intention of *removing* redundancy, complexity, and distraction, so as to leave (ideally) only essential elements. Proponents of this approach seek to relieve the observer of potentially confusing stimuli



Sensory Feedback and Remote Control of Machines in Mining and Extraterrestrial Environments

but run the risk of an overly-impooverished display. Which is the better of these approaches is not always evident.

Vision, touch and audition normally work in parallel. They complement each other when they all agree with what is being perceived. When they disagree (suffer incongruence) confusion can result with vision often reported to be the dominant sense.

Movement (kinesthetic) information, and tactile (cutaneous) information such as vibration and texture, complement each other when the skin-sensed texture of an orange, for example, agrees with the shape and pliability of it, as sensed by the position of fingers grasping it. However, until recently, it has been difficult or impossible to test these processes separately and determine the relative importance of movement and skin sensations, and whether actively exploring something leads to better perception than being guided passively. These issues are important to designers who want to know how best to provide haptic feedback, especially when enhanced feedback, and choice of input, is possible.

Recently, at the Bionics and Cognitive Science Centre of Monash University, it has been possible to shed light on how to optimise haptic displays. We used a device we call the Tactile Display System, or TDS (Richardson & Symmons, 2000) with some surprising results. This device can record and play back exploratory movements with great precision, and it allows the contribution of cutaneous inputs to be individually assessed. We have found that as much information about a raised line drawing can be conveyed through a stationary fingertip (which the drawing is simply passed across), as is conveyed by moving a person's finger along a pathway corresponding to the drawing's outline (Richardson, Symmons & Wuillemin, 2004). The former condition involved touch (cutaneous information) alone, and the latter, kinesthetic information alone. Previous research predicted that movement (kinaesthetic) cues would be much more important than touch. In other experiments, we have found that being passively guided around a raised line drawing can lead to better perception than being allowed to freely (actively) explore (Symmons, Richardson & Wuillemin). These findings have implications for training in simulated environments such as those used for teleremote surgical training, as well as machinery operation. We are currently extending this research from two-dimensional to three-dimensional explorations.

CONCLUSIONS

There are sound practical and theoretical reasons to explore the benefits of multisensory feedback to operators of remote machinery, whether for mining, bomb disposal, or extraterrestrial exploration. The practical reasons have to do with cost-effectiveness. Performance should improve along with longer shifts and lowered risks to operator and machine. The theoretical reasons include an increase in knowledge that can be applied to training simulators, teleremote surgery, and artificial intelligence. Until robots have sensory/perceptual and cognitive processes rivaling those of humans, remote control is a viable option that can be better than

proximal control, and can contribute to the design of fully automated (thinking) machines.

Note: The Bionics and Cognitive Science Centre at Monash University, and WMC Resources Pty Ltd, are currently investigating the possibility of retrofitting a remotely controlled rockbreaker with multisensory feedback capacity. ■

References

- Bradshaw, M.F., Elliot, K.M., Watt, S.J., Davies, I.R. (2002). Do observers exploit binocular disparity information in motor tasks within dynamic telepresence environments? *Proc. SPIE Vol. 4660, p.331-342. Stereoscopic Displays and Virtual Reality Systems IX*, Andrew J. Woods; John O. Merritt; Stephen A. Benton; Mark T. Bolas; Eds.
- Draper, J.V., Handel S., and Hood C.C. (1991). "Three Experiments with Stereoscopic Television: When it Works and Why", *Proceedings of IEEE International Conference on Systems, Man and Cybernetics, Charlottesville, Virginia*.
- Goldstein, B.E. (1999). *Sensation and perception*, 5th ed., Pacific Grove: Brooks Cole.
- Heller, M.A. (1983). Haptic dominance in form perception with blurred vision. *Perception*, 122, 607-613
- Matti, G. (2002). Is audio useful in immersive visualization? *Proc. SPIE Vol. 4660, p.411-421. Stereoscopic Displays and Virtual Reality Systems IX*, Andrew J. Woods; John O. Merritt; Stephen A. Benton, Mark T. Bolas; Eds.
- Palmer, S. (2002). *Vision Sciences*. Massachusetts: Bradford Books, MIT Press.
- Rastogi, A., Milgram, P., & Drascic, D. (1996). Telerobotic control with stereoscopic augmented reality. *SPIE Vol. 2635, Stereoscopic Displays and Virtual Reality Systems III*. Mark Bolas; Scott Fisher; & John Merritt; Eds., pp.115-122.
- Rauterberg, M. (1999). New directions in User-System Interaction: augmented reality, ubiquitous and mobile computing. *IEEE Proceedings Symposium Human Interfacing, May 20*, pp. 105-133.
- Richardson, B.L., Symmons, M., & Accardi, R. (2000). The TDS: A new device for comparing active and passive-guided touch. *IEEE Transactions on Rehabilitation Engineering*, 8, 414-417.
- Richardson, B.L., Symmons, M., Wuillemin, D.B. (2004). The relative importance of cutaneous and kinesthetic cues in raised line drawing exploration. *Proceedings of the Touch, Blindness and Neuroscience Conference*, Madrid, Oct 2002.
- Roberts, J.W., Slattery, O.T., Swope, B., Volker, M., & Comstock, T. (2002). Small-scale tactile graphics for virtual reality systems. *Proc. SPIE Vol. 4660, pp. 422-429. Stereoscopic Displays and Virtual Reality Systems IX*, Andrew J. Woods; John O. Merritt; Stephen A. Benton; Mark T. Bolas; Eds.
- Sarter, N. (1998). Turning Automation Into A Teamplayer: The Development of Multisensory and Graded Feedback For Highly Automated (Flight Deck) Systems, Willard Airport Aviation Research Lab One Airport Road - Bldg. Q5.Savoy IL61874
<http://nsf-workshop.engr.ucf.edu/papers/Sarter.asp>
- Schmit, A., Grasník, A. (2002). Multiviewpoint autostereoscopic displays from 4D-Vision GmbH. *Proc. SPIE Vol. 4660*, pp. 212-221. *Stereoscopic Displays and Virtual Reality Systems IX*, Andrew J. Woods; John O. Merritt; Stephen A. Benton; Mark T. Bolas; Eds.
- Shiffman, H.R. (1996). *Sensation and perception* (4th ed.). New York: Wiley.
- Woods, A. (2003). Seeing in depth at depth. *Newsletter of the Centre for Marine Science & Technology*. September, 2003.