Remote Reality via Omnidirectional Imaging

Terry Boult tboult@eecs.lehigh.edu
Vision and Software Technology Lab, EECS Department Lehigh University

Abstract

For a myriad of military situations, video imagery provides an important view into a remote location. These situations range from remote vehicle operation, to mission rehearsal, to troop training, to perimeter security. In some situations, the system needs a long focal length to see distant objects. In many others, however, it needs to see a wide area and often benefits from the ability to view in different directions.

DARPA and ONR funding has led to the development of new technologies that may radically alter the way we handle remote viewing of wide areas. By combining omni-directional imaging and a head-tracked HMD,*Remote Reality* provides an immersive window into the remote environment. The system provides a 30 frame-per-second (fps) video stream of the remote site in whatever "direction" the user looks. Unlike pan-tilt units, it supports multiple simultaneous users looking in different directions, and is capable of updating its viewing directions with only a 60 to 30 ms delay, i.e. 15-30fps . This paper describes some of the details of the system and its applications.

1 The Remote Reality System

The main components of the Remote Reality system are the omni-directional camera, video recording/transmission systems, and a head-mounted-display (HMD). Remote Reality is a direct application of omni-directional imaging which will be discussed at end of the paper. First let look at the system, its interfaces, and some applications.

The first prototype immersive system uses COTS parts and strove to minimize cost while maintaining acceptable quality. Our current data collection system was approximately $4K (+$1K for underwater) and the computing/HMD play-back system was about $3K. The system uses a 233Mhz K6 CPU (running Linux) and a $300 video capture card. The system computes biocular 320x240 30 fps NTSC video. This resolution is reasonably matched to the low-end HMD used, Virtual IO's I-Glasses. The HMD's built-in head tracker provides yaw,pitch an roll, with updates to the viewing direction at 15-30fps. With a better head tracker (e.g., Intersense IS-300) and a 300Mhz CPU we can insure consistent 30fps update of both viewpoint and video data. Better HMD's are also commercially available, at costs ranging from $2K to $10K, for low to medium volume usage units, and $20K for very rugged high-volume usage.

While a number of paracamera models are commercially available (see www.cyclovision.com) most of our Remote Reality systems use our own smaller custom designs directly incorporating camcorders rather than camera, e.g. see figures 1–2. The development of the underwater cameras and vehicle cameras involved solving both optical and mechanical design problems. At IUW98 we will demonstrate Remote Reality using video from underwater, ground, and possibly helicopter systems.

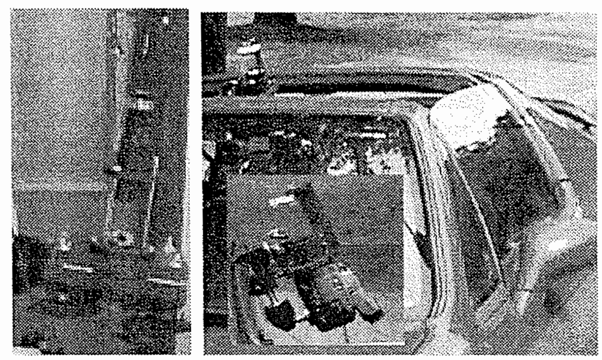

Figure 1: Two generations of omnicamera car mounts.

Figure 1 shows some custom car mounts for omnicameras. The first generation vehicle mounts, which attached to the car windshield or roof via suction-cups and straps, used the Cyclovision paracameras and a separate video recorder inside the vehicle. The second generation systems use our custom design with optical folding and integrated cam-corder which puts the camera inside the vehicle. To use it one simply pops the mirrors above a sun roof. In both designs cases, damping vehicular vibrations are an issue; for hand-held usage of the folding design, a gyro-stabilizers is recommended.) For furtive data collection we are now developing a mount, mounted inside a "pizza sign", that has room for multiple cameras with with integrated gyro-stabilizers.

There are two current interfaces, each intended for different applications:

- *immersive*: giving the user the impression they are at the remote application.
- *augmentive*: enhancing the users situational awareness.

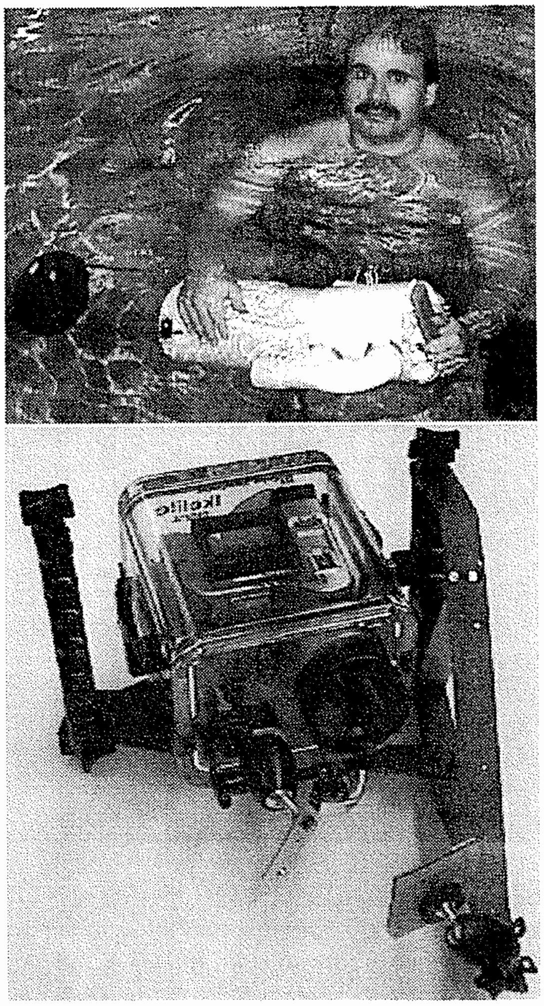

Figure 2: First and second generation underwater Omnicameras. First generation was a low-cost home built unit intended for proof of concept and initial empirical system evaluation. The second generation uses optical path folding for more compact design and so operator can stay behind camera.

1.1 Immersive Remote Reality

The first interface we developed, the immersive interface, uses a a biocular HMD with a head tracker. The head tracker provides orientation information which is used to determine the viewing direction for the the unwarping map. As the HMD turns, or if the user requests a software "zoom," the virtual viewpoint is stationary; only the virtual "imaging array" is moved. This provides a very smooth and natural visual change. However, maintaining immersion depends on a truly responsive system. Making the system fast took a few, but straight forward tricks including: fixed point math for most computations and table lookup for expensive operations.

In the immersive interface the user sees nothing but the video. Applications where the user is active in their observation but passive with respect to motion control benifit the most from this interface. The immersive interface is particularly effective in applications intended to acquaint the user with a remote environment, it appears as if they are a passenger at the remote location. Education, training, and mission rehearsal are obvious candidates for recorded Remote Reality. A few esoteric applications include: cataloging the state/contents of complex facilities such as ships or pipe complexes; security surveys of a route or building.

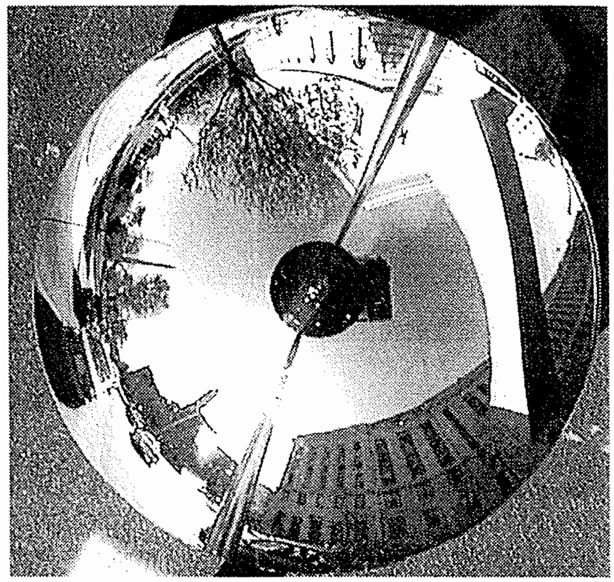

Figure 3: An Omnidirectional (Paralmage) taken in Washington DC. The Washington monument in the left of the image.

1.2 Augmentive Remote Reality

The second Remote Reality interface, Augmentive Remote Reality, is being developed for applications where the user needs to augment their reality, rather than supplant it. The driving applications here include remote vehicle operation and urban police actions. Both ground and helicopter-based systems are being developed/tested.

Immersing oneself in a remote video stream is generally not sufficient for vehicle operation. Although a head-tracking interface is natural for vehicle operation, the user needs additional information such as vehicle position, direction and speed. Like existing systems, we will integrate GPS (or DGPS) and inertial navigation. The pilot in a safe location could use the biocular HMD with head tracking for setting view direction leaving their hands free to operate the vehicle. If they drive in the field, a monocular head-tracked system can be used. Informal observations suggest that for simple environments a pilot spend most of their time facing directly ahead, but as the environment becomes more complex and the desired path includes many turns, the pilot increasingly uses their freedom of viewing direction. Preparations are underway for formal evaluations. Other than the speed of response, using Remote Reality for a solo pilot is not significantly different than having a remote pan/tilt unit. The difference becomes significant when team members also need to significant view non-navigational features within the environment.

In urban maneuvers, vehicles can be piloted from a relatively safe location, but someone usually needs to be following it for clearing/securing activities. The vehicles can transmit (encrypted) omni-directional video while team members use augmenting Remote Reality to look for potential threats around the vehicle's location. Unlike what could be done with a pan-tilt system, *the team members can simultaneously look in different directions*. A soldier can litterly watch his/her own back. Additionally no team member needs to transmit to the vehicle to control the pan/tilt viewing direction; the forward team can all be radio silent.

For the wearable augmented Remote Reality system we have been examining interfaces ranging from the raw omni-directional video with possible overlays (minimal to no computer needed) to displays showing wide-view panoramas or smaller perspective views. View control for the wearable system can be a mixture of head-tracking and/or track-ball control. We are also developing a semi-automatic motion-target tracking based on our tracking work[Boult *et al.*-1998a]. The user interface is expect to continue to evolve significantly as we gain experience and feedback from potential users and system testing.

In recent years, [Nayar-1997b, Peri and Nayar-1997], Dr. Nayar set the stage for Remote Reality and demonstrated software unwarping paraimages in real time on standard PC hardware. Cyclovision now sells both cameras and mouse/joystick controlled WindowsNT software. The advances needed for Remote Reality were mostly systems integration and interface issues: algorithm improvements to provide 15-30fps view updates and the integration with the head tracker and HMD. We also developed code to do 30fps warping including efficient multiplicationless 4-case factor-of-2 bilinear-based interpolation done directly in RGB555 color space.

Unfortunately, none of commercially available wearable computers had the video I/O bandwidth and resolution necessary for the 640x480 30fps video processing. We have built (using COTS parts), a wearable version of the of the augmented Remote Reality system and expect to demonstrate it at IUW.

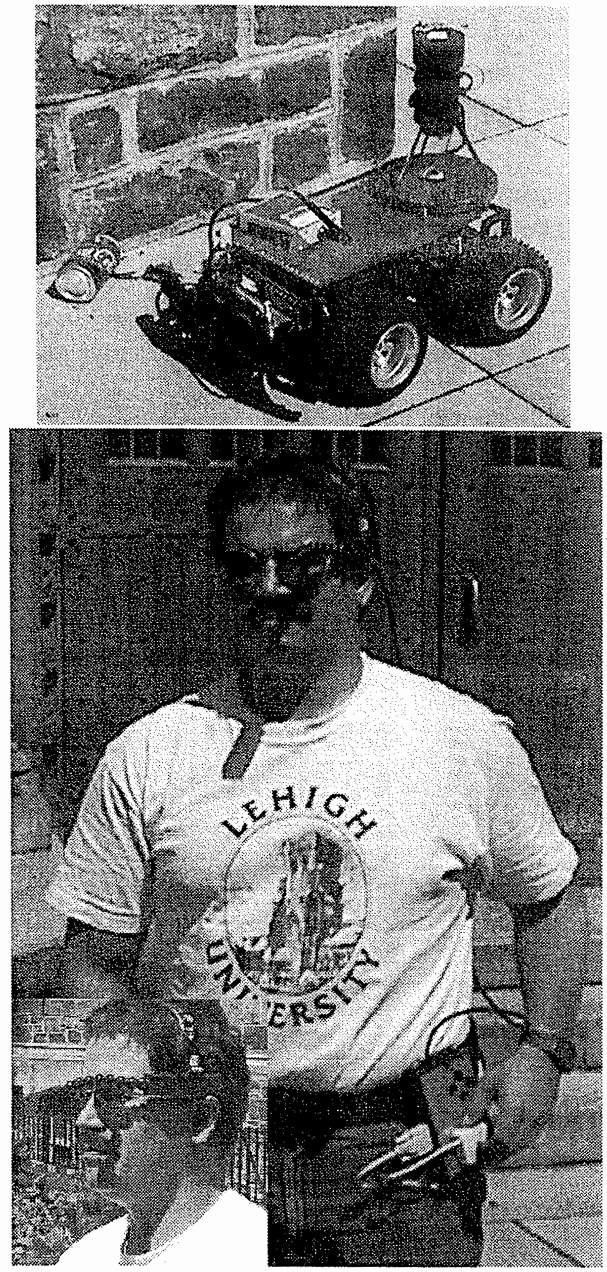

Figure 4: Omni-directional vehicle and and augmented wearable Remote Reality.

2 Omni-Directional Imaging

While Remote Reality systems could be built with a multitude of cameras at the remote location, central in its design was the omni-directional camera designed by Shree Nayar [Nayar-1997a]. This camera directly captures a full hemisphere (or more) while maintaining a single perspective viewpoint allowing it to be used for full motion omni-direcional video. Furthermore, placing two of these paracamera systems back-to-back allows a true viewing sphere, i.e. 360 x 360 viewing. Unlike fish-eye lenses, each image in the paracamera system can be processed to generate *geometrically correct* perspective images in any direction within the viewing hemisphere.

The paracamera's omni-directional imager combines an orthographic lens and a parabolic mirror with the axis of the parabolic mirror parallel to the optic axis of the lens systems. The orthographic lens results in the entering rays being parallel. Rays parallel to the axis reflect off a parabolic surface at an angle such that they virtually intersect at the focus of the parabolic surface. Thus the focus of the parabolid is the single "virtual" viewpoint critical to Remote Reality systems. A single viewpoint allows for consistent interpretation of the world with a very smooth transition as the user changes their viewing direction. While there are other systems with large or even hemispheric fields of view, as show in [Nayar and Baker-1997], fish-eye lens and hemispherical mirrors do *not* satisfy the single viewpoint constraint.

To generate a proper perspective image from the para-image, consider an "imaging array" in the desired viewing direction. For each pixel, logically cast rays through the focus of the parabolic surface and intersect them with captured image. Thus for each pixel in the perspective image we can obtain its corresponding position in the orginal image. The resulting spatially varying resampling can be very efficiently implemented using spatial lookup tables. Because resolution is a serious issue when a viewing hemisphere is packed into single image, we have implemented the efficient bilinear-like mentioned earlier. To speed up the warping table computation for new viewpoints, the computations are "incremental", and use all fixed-point mathematics and table lookup for the complex components.

Because omni-directional imaging compresses a viewing hemisphere into a small image, maintaining resolution and captured image quality is quite important, and takes careful design. While the process scales to any size imager, the current systems use NTSC (640x480) or PAL (756x568) cameras. The spatial resolution along the horizon is $\frac{240\text{pixels}*2*\pi}{360\text{degrees}} = 4.2\frac{\text{pixels}}{\text{degrees}}$ (5.1 for PAL). Note the "spatial resolution" of the image is not uniform. While it may seem counter intuitive, the spatial resolution of the omni-directional images is *greatest* along the horizon, just where objects are most distant. Cyclovision now sells a 1Kx1K still camera and we have built a 1Kx1K @ 5fps system. Higher resolution/speed systems are being developed, although they will be considerably more expensive than those based on consumer cameras.

Some graphics/VR-oriented researchers who have not experienced Remote Reality first hand, may consider the 320x240 15-bit resolution inadequate. The system, however, has been demonstrated to a large number of people with very positive feedback, see [Boult *et al.*-1998b], [Boult-1998b] and [Boult-1998a]. While the resolution is far from that of high end graphics systems, the naturalness of objects, fluidity of motion and the complex/subtle textures (even at low-resolution) of the video seem to make up for the pixel loss. The interface and system continue to evolve but even at the current state Remote Reality offers significant advantages.

[Boult *et al.*, 1998a] T. Boult, Ali Erkin, Peter Lewis, Ross Micheals, Chris Power, Chen Qian, and Weihong Yin. Frame-rate multi-body tracking for surveillance. In *Proc. of the DARPA IUW*, 1998. These proceedings.

[Boult *et al.*, 1998b] T. Boult, Chen Qian, Weihong Yin, Ali Erkin, Peter Lewis, Chris Power, and Ross Micheals. Applications of omnidirectional imaging: Multi-body tracking and remote reality. In *Proc. of the IEEE Workshop on Computer Vision Applications*, October 1998.

[Boult, 1998a] T. Boult. Remote reality. In *Proc. of ACM SIGGRAPH 1998*, 1998. Technical Sketch.

[Boult, 1998b] T. Boult. Remote reality demonstration. In *Proc. of IEEE CVPR*, 1998. Technical Demonstration.

[Boult, 1998c] T. Boult. Remote reality via omnidirectional imaging. In *Proc. of the DARPA IUW*, 1998. These proceedings.

[Nayar and Baker, 1997] S. K. Nayar and S. Baker. Complete Class of Catadioptric Cameras. *Proc. of DARPA Image Understanding Workshop*, May 1997.

[Nayar, 1997a] S.K. Nayar. Catadioptric omnidirectional camera. In *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition*, pages 482–488, June 1997.

[Nayar, 1997b] S.K. Nayar. Omnidirectional video camera. In *Proceedings of the 1997 DARPA Image Understanding Workshop*, May 1997.

[Peri and Nayar, 1997] V. Peri and S.K. Nayar. Generation of perspective and panoramic video from omnidirectional video. In *Proceedings of the 1997 DARPA Image Understanding Workshop*, New Orleans, May 1997.