Panospheric Video for Robotic Telexploration

John R. Murphy

CMU-RI-TR-98-10

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Robotics

The Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

May 1998

This research was supported in part by NASA grant NAGW-1175. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of NASA or the U.S. Government.

Carnegie Mellon

Robotics

Thesis

Panospheric Video for Robotic Telexploration

John Murphy

Submitted in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy
in the field of Robotics

ACCEPTED:

William L. Whittaker	Thesis Committee Chair	8 May 1998 Date
Matthew T. Mason	Program Chair	1 June 1998 Date
Raj Reddy	Dean	6/5/98 Date

APPROVED:

Paul Christiano	Provost	30 June 1998 Date

Abstract

Teleoperation using cameras and monitors fails to achieve the rich and natural perceptions available during direct hands-on operation. Optimally, a remote telepresence reproduces visual sensations that enable equivalent understanding and interaction as direct operation. Through immersive video acquisition and display, the situational awareness of a remote operator is brought closer to that of direct operation. Early teleoperation research considered the value of wide fields of view. However, utilization of wide fields of view meant sacrificing resolution and presenting distorted imagery to the viewer. With recent advances in technology, these problems are no longer barriers to hyper-wide field of view video.

Acquisition and display of video conveying the entire surroundings of a teleoperated mobile robot requires innovation of hardware and software. The raw image from a panospheric camera is unfamiliar for direct viewing by a human operator. Through software transformation, an image is created that is visually identical to images produced by traditional field of view cameras. The image stream can be formatted to mesh with any display, and is displayed on both a flat-screen monitor and a domed immersive theatre.

This research explores the value of panospheric video to telexplorers who navigate remote mobile robots through unknown terrain. A substantial component of the research is the design and development and evaluation of a panospheric sensor and methods of displaying the video to the telexplorer. Different modes of presenting the video to operators of a remote robot are evaluated based on the success and techniques of the operator while attempting to react to untraversable terrain.

Utilization of panospheric video for robotic telexploration is an appropriate solution, even though current technologies fall short of achieving the characteristics of human vision. In its initial deployment, panospheric technology shows significant value over traditional methods of image acquisition. Utilizing panospheric video, over 100 novice and experienced operators combined to accomplish unprecedented teleoperation production.

Acknowledgment

Many have contributed to this work. This thesis wouldn't exist without an initial spark from Steve Bogner, PVSI and DRES. Thanks go to Dave Lavery for insights, and Eric Krotkov, Hagen Schempf and Butler Hine for always being available, interested, and constructive whenever I sought them. Blessings to Eric Rollins and the Lunar Rover Initiative, for incarnating bold ideas and tolerating devil's-advocate pessimism.

Just as Art Charlesworth and Greg Morrisett were my inspiration at Richmond that brought me to CMU, Red Whittaker and Lalit Katragadda are the friends who have stood by my side during these years and at whose side I shall create the future.

I thank Michael Nechyba, Martin C. Martin, Cassandra Briggs, and Ann Musselman who took away the monotony of grad school existence with educational life experiences such as multiple Big-Screen college football viewing, arguments over movies and politics, skiing on the purported 38inch base, sip-a-frag Doom© marathons, birthday party crashing at night-club's, the development of the "lunch" as a currency, phone conversations about the crazy old times and the crazier new ones, concert attendance, spring breaks and beach weeks that defied planning.

I remember with fondness Mark Ollis, Barry Brumitt, and Deepak Bapna for pre-5pm foosball, CDs on repeat, ringer/answering machine volume wars, shared computer trouble, and other worthwhile distractions as only officemates in the latter years of a Ph.D. process can provide. Thanks too to the RI with all its wonderful people (including Dot Marsh and Marce Zaragoza). Unique thanks to Highway 522, the Appalachian mountains, 3rd Street Diner, and Topsail Island for always being only hours away.

The debt I owe to my parents and sister and grandparents and other family & friends for their support from 1st to 23rd grade is so vast I can only say: "Thank you. I love you."

Perhaps greatest of all, in exchange for her unwavering voice of reason even during phone calls at 4am, drive and dedication that ran both of us at times, open ears whenever I needed to talk, and an open heart when there wasn't anything to say, I extend love and "break-a-leg" wishes, regardless of where her career leads, to Amanda, my best friend.

I also thank those that made the Stateside Operations of the Atacama Desert Trek happen so successfully, including Paul Oles, Tom Flaherty and the CSC staff; SimLab, Spitz Inc. and Bally Design; Michael Sims, David Wettergreen, Kurt Schwehr, Maria Bualat and everyone else at NASA Ames Research Center's Intelligent Mechanisms Group who have turned telescience into an art form; Dr. Geb Thomas and his cohorts at the University of Iowa GROK lab who made the telepresent interface a reality; the Members of the Center for Metahuman Exploration who brazenly undertook the creation and delivery of RoverTV; and Anne Watzman, Leanne Crosbie, and all the members of the media who provided their time and interest in bringing the Atacama Desert Trek and telexploration to the local community and world at large. Thanks too to RedZone Robotics for their efforts to further mature panospheric imaging through SBIR work.

With their gracious contributions of hardware and software, I am endebted to John Coghill and Kathy Kline at DALSA Inc., Leonard Mygatt at Summus Ltd. the staff at Vision1 and Bitflow, Stuart Henly at MIL Inc., and the great guys at Bernie's Camera on Pittsburgh's north-side who probably remember me to this day as a raving loony.

Without the brilliant handling of both the public and Nomad during the course of the Trek by Lauren Garcia, Lisa Ochs, and Emily Weiss this thesis would likely never have been accomplished. And lastly, I want to thank all those who took time out of their 1997 summer visit to Carnegie Science Center and willingly took the controls of a robot a continent away and unwittingly became part of this thesis:

Joey Weitz, Penina Weitz, David Saforo, Michele Eaton, David Nedde, Dan Foygel, Bill Brooks, Nancy Brooks, John Hancock, Andy Mor, Daniel Huber, Alexi Morrisey, Peter Coppin, Phil Berg, Jen Sever, Michael Tung, Steve Patterson, Mike Patterson, Tim Patterson, Amanda Dick, David Mikesell, Mary Mikesell, William Mikesell, Steffi Yessler, Caitlin Lee, Davie V., Joshua Axzinger, Matt Slafha, Will McLaughlin, Tom Higgs, Bill Byrne, J.R. Richard, Clayton Holt, Danielle Fischer, Evan, Scott, Ben Brooks, Sinan, Daniel Nickeson, Jeff Nickeson, David McGaston, Jen Madaffer, Dare DuPlain, Victor Fanks, Zachary Fanks, Christopher Dowdy, Rebecca Idell, Larry Garcia, Laura Mitchell, T.J. Hogue, David Pini, Jessica Pini, Maryann Pini, Mike Pini, Aaron Matthews, Chris Taylor, Brandon, Cody Hitt, Jake Hitt, Andrew Moore, Dave Hardle (the telexplorer at the controls when Nomad accomplished the 200km milestone), Abe Anderson, Alex Sycara-Lewis, Lisa Kouvolo, Stephen Shetler, Pauline Shetler, Jay Gene, Maloomo Thompson, Anthony Brine, Avid Melle, Jada Hale, Rob Logan, Mark Hudson, Marissa Gurkey, Duane Ramseur, Tishla Jones, Danielle Williams, Mike Zerega, Me Chen, Taine Ogasaware, and 27 "anonymous" drivers.

Thank you all.

Contents

Chapter 1

Immersive Teleoperation

The cameras and monitors of traditional teleoperation do not provide the rich and natural perceptions available during direct hands-on operation. A remote telepresence would ideally reproduce visual sensations that enable equivalent understanding and interaction akin to direct operation. While the future promises autonomous robots and altered operator roles, teleoperation remains the technology of choice for tasks which focus on extending the human presence and intellect, such as exploration. As robotic roles become more challenging, teleoperation must become easier to learn, more productive and safer.

Through hyper-wide field of view video acquisition and immersive visualization, the situational awareness of a remote operator is brought closer to that of direct operation. Although replicating human vision is not a capability of today's technology, it is possible to create a teleoperation sensor that achieves usable panospheric imaging (simultaneous views in all directions). Video delivered from such a sensor can be displayed in many ways from fully immersive to traditional perspective views or in hybrid formats. This research explores the effects of panospheric video on telexploration and examines numerous visualization options.

Motivation

Popularized robots are personable, do-all, be-all, sentient machines as fantasized in "The Jetsons", "Dr. Who", "Lost in Space", and "Star Wars". In many works of science fiction, humans interact on a nearly even level with robots. *Rosie*, the Jetson's house-robot, performs all the mundane tasks about their home, from cooking to cleaning, watching the kids, and even may be consulted on a family matter. The two "droids" *R2D2* and *C3PO* in Star Wars are contrived to traverse deserts and corridors, translate languages, and perform complex calculations [Lucasfilm97].

Robotics research has brought about fantastic capabilities, but the carrot held by the science fiction community remains beyond reach of today's robots. The modern robot that is most prolific is the manipulator arm, which has successfully taken over many assembly line tasks that require no skill other than accurate repetition. In the research communities, robots that walk, talk, navigate, and manipulate have all been demonstrated, but a significant gap remains between science fact and fiction.

For tasks that are extensions of the human spirit, such as exploration, teleoperation is more appropriate than autonomous operations. While robotic cognition and perception skills are being honed in the laboratory, "dumb" robots can be placed in remote locations and utilized as tools by a remote human operator. These robots make few judgements about the realm they explore, and provide sensory feedback of the remote location. The operator pieces together an awareness for the robot's environment from the various sensor data, and can use their knowledge and reasoning skills to make deductions and conclusions about the area being explored.

Situational Awareness

Much of an operator's efforts during exploratory teleoperation are spent avoiding or extricating from trouble. When commanding a robot to move about a remote environment, the operator needs to consider the limitations of the robot and correctly interpret a remote terrain. The mobility of robots is below humans in speed and agility and ability to change mode (walking to running to crawling to leaping, etc.). When approaching difficult terrain, the operator must recognize it as problematical for the robot, and begin to take action.

Human cognition distinguishes us from the animal kingdom. However, valid cognition is reliant on valid perception of the environment for making successful decisions. Thinking about how to handle a stretch of difficult terrain, or planning a way to escape from a problematic situation is an arena where humans have the advantage over a robot. However, unless the operator has information about the remote situation through the perception sensors of the robot, cognitive skills are not relevant.

80% of the sensory input received by humans is visual, and robot operators are most comfortable when visual information about the remote environment is available. When a human interacts with the environment, the senses of sight and sound and touch are used much more often than taste and smell. And compared to sight, sound and touch are used sparingly when interpreting an unknown environment [Almaula86].

Both the foveal and peripheral components of human vision play a large role in ensuring deft operations, but peripheral imagery is more important for maintaining heightened awareness. The foveal component of human vision provides the resolution to recognize features. The peripheral component allows for the features to be associated with other features to form objects and scenes. In order to safely operate in an unknown environment, the foveal resolution required depends upon the size of the smallest object that must be identified (e.g. even 20/40 vision passes most driver-license exams). How much of the environment must be known to operate without risking collision or damage depends on the shape and characteristics of the "vehicle", whether it be human or robot (e.g. opening the car door to check closeness to the curb when parking). The peripheral component of human vision provides the awareness needed to move through an environment while keeping track of all appendages, and hence navigate safely.

In an unknown environment, seeing more views of an object provides more knowledge. Since virtually all objects in a remote environment are stationary, acquiring multiple views requires being in multiple locations. A mobile robot can maneuver about an object and gather these views, and then move on to a new object. Thus mobile robots both require and provide situational awareness during operations.

Single images at multiple locations are sufficient to provide understanding of simple shapes (e.g. top/front/side views of objects in blueprints). However, natural objects and terrains are often less symmetric and structured, and many views are required for full understanding. Still-frames require the operator to extrapolate; video continuously updates the scene. Without video, the operator is denied current feedback crucial in dynamic environments, and must make worst-case assumptions when attempting to proceed (e.g. never commanding the robot to move into an area that cannot be seen from the current location). The Lunokhod and Sojourner planetary rovers were limited to still-frames, and exploratory progress was tedious and slow.

As the purpose of telexploration is to gather information about something of which the operator is unfamiliar, the goal is to provide a telepresence experience that enables heightened situational awareness. With accurate awareness of the remote environment, the operator can act in a manner more commensurate with actually being at the remote location. Visual information contributes substantially to achieving awareness, and the peripheral, as well as the foveal, component of imaging is desired.

Current Imaging Technology

Robotics has capitalized on improvements in computer processing and memory capabilities, but imaging technology has somewhat been neglected. Advances in high-speed computing enable algorithms that once had to be ignored due to complexity or time consumption, especially in planning and perception. Imaging technology has also improved, with commercial hand-held digital cameras and image-processing software. However, most designers still consider imagery to be something that is directly passed from the camera to the operator with little alteration.

Most teleoperation sensors are cameras with around 50degree fields of view. These fields of view are used due to their ability to provide undistorted images that the operator can view directly and yet are not too restrictive. With the advances in camera resolution, these cameras are becoming closer to imitating the human foveal vision component [Kodak97].

To achieve the affect of the peripheral component, immersive imagery is acquired. The two traditional methods for this acquisition are camera arrays or pan/tilt mechanisms, both of which have disadvantages for robotic telexploration. Multiple cameras increase the complexity of the system and require objects to be tracked over multiple images or the imagery to be stitched together. Pan/tilt cameras can only view one object at a time and require actuation, making generation of immersive video streams cumbersome.

Prior work in teleoperation suggested the advantages of wide field of view, but dwelled on the difficulties in acquiring and interpreting the imagery. The number of cameras required in an immersive imaging array is directly proportional to the field of view of the cameras. Wide field of view imaging made more sense, as it decreases the complexity of the camera array, but the added distortion in the images made teleoperation using the video less effective. In addition, as the field of view increased the resolution at each point decreased, and so the video became less useful [Glumm92].

An innovative solution which acquires views in all directions simultaneously is suggested, and components required for "panospheric" technology are now available. With the advances in imaging technology, immersive video can be obtained from wide field of view imagery with usable resolution and distortion that can be processed away in software. Lensing that combines refractive and reflective optics can gather full immersive imagery without requiring multiple cameras.

Current Display Technology

Recreating telepresence requires a suitable display media. To give the impression of actually being at the remote location, the ideal imagery display would replace the local surroundings with the remote views. It would also provide sufficient resolution at the foveal viewpoint, while filling the peripheral view.

Monitors remain the principal display medium for teleoperation imagery due to their low cost and common availability. Arrays of monitors are cumbersome and the seams are often visible. Stadium jumbo-tron's capitalize on distance to hide the seams between monitors, but currently are utilized to display the standard field of view signal normally delivered to a single monitor. Monitors or flat LCD screens make sense when the imagery being returned is a non-immersive perspective projection. Seaming together banks of monitors converts what the eyes are viewing as a sphere into a polygonal surface, which adds parallax effects. In addition, it is difficult to position enough monitors to fill the peripheral view completely.

Head-mounted displays are improving dramatically, but lean toward reproduction of foveal views instead of providing simultaneous WFOV imagery. Head-mounted displays are really not much different than a pair of monitors hung very close to each eye. There is no attempt to allow display for the peripheral regions. Also, the technology to adjust the imagery as the head moves is not sufficient to mimic natural viewing, and can often make users nauseated. However, the head-mounted display is a strong candidate for the eventual telepresence solution.

Domed or curved-screen displays can achieve immersive FOV, but are utilized mainly for simulated environments or post-processed films. Planetariums, OmniMax theatres, and high-end simulators utilize domed displays. The user is successfully immersed in the remote environment. Imagery for these displays is mostly created via graphics and simulation, and must be formatted to match the particulars of the display in terms of projector number and position and view. However, using panospheric technology, immersive displays can be filled immediately from on-board a robot, enabling their use in teleoperation.

Thesis Statement

Robotics technology has reached the point where difficult tasks can be accomplished in even the most hazardous of environments. As tools, robots allow for enhanced accuracy, precision, strength, and endurance. From the factory floor to volcanic craters, robots have demonstrated their ability to perform. While autonomous operations or science-fiction robots are not broadly viable, robots do extend the human presence deep under the sea and on other planets.

Exploration of remote locations is a common thread in human history, but higher value placed on human life and higher cost to find remote locations presents a need for telexploration. In centuries past, it was accepted that those sailing beyond the horizon might never return home, but the lure of what lay beyond was too strong to ignore. In current times, travel between continents is commonplace, and the view of a far-off-land can be obtained through the internet. These factors reduce the desire to perform life-risking exploration, but feed the desire to explore the world and beyond. The solution which satisfies these desires is telexploration.

Providing video feedback increases the telexplorer's sense of "being-there" over still-frame images or no imagery at all. Without returned imagery, the telexplorer must use other data to reconstruct the environment and translate it into something that is easily understood. Terrain maps and superimposed imagery or panoramic stills provide some feeling for the remote location, but can not accurately portray dynamic environments, or give the telexplorer all the views needed to evaluate and adjust to terrain not suitable for the robot.

This research investigates the thesis assertion:

> that video conveying the entire surroundings of a teleoperated mobile robot enhances situational awareness needed to react to untraversable terrain while exploring.

To perform the work, innovation was required to acquire and display "video conveying the entire surroundings of a teleoperated mobile robot". As a result, a substantial component of the research is the design and development and evaluation of a panospheric sensor and methods of displaying the video to the telexplorer.

Scope and Approach

The areas of teleoperation and situational awareness are general enough to enable substantial differences of interpretation among readers. To help focus the reader, as well as the research, the scope of the work is limited as follows:

- **wheeled mobile robot**. The robot genre for which the sensor was designed, and upon which its evaluation was performed were "mobile robots" rather than fixed systems such as manipulators. More specifically, the robots were wheeled vehicles that needed to be commanded much like an automobile (with direction and speed). Most of the telexplorers who participated in the research were familiar with how cars are driven, but were too young to have their license and actual driving experience. This was considered valuable as the robot did not have traditional ackerman steering, and so the telexplorers were not influenced by related automobile domain knowledge.
- **unknown environment**. To ensure that no prior knowledge existed to aid in achieving awareness, an unknown environment was utilized. None of those who utilized the video to "telexplore" had ever been to the Atacama Desert before, or had seen any pictures of the robot in the actual terrain that might give cues to understanding scale beyond what

was provided in the video. It also meant that perceptions were not clouded by the resolution or color of the video as being "not what i'm used to", as they had no way of knowing what the environment would look like to their own eyes.

- **real world**. The environment of operations was the real world, not a simulated or laboratory environment. Terrain that was "untraversable" occurred naturally, and the surroundings of the robot was a natural setting. Modeling and rendering of natural terrain and objects at video speeds is a separate research agenda, and using the real world enabled the research without resorting to simplifications required in simulation. The downside was that every telexplorer had a different experience and so quantitative results comparing different operators on set tasks were not possible. It also meant that there was terrain encountered that could destroy/disable the robot and telexplorers could not learn the capabilities of the robot from "trial and error".
- **remote operations**. The teleoperation modality used was fully remote "inside-out" operations rather than direct control or over-the-shoulder "outside-in" operations. The only access the telexplorers had to the unknown environment was through the data returned via the robot. There were no off-board views, or descriptions of the environment provided by those who could see the robot directly. Telemetry data available beyond the imagery was roll and pitch inclination and vehicle speed and compass direction.
- **concept demonstration**. The emphasis of the sensor development and video display was not on refinement, packaging, or perfection, but on adequate and functional demonstration of the concept. There were known weaknesses with the sensor and the display such as color balance, resolution, glare, blooming. Yet, as all these features were detrimental to the telexplorers, any positive results from the research remain valid, and would possibly be enhanced by a refined system.
- **mono viewing**. Humans garner substantial awareness about the environment from natural stereoscopic viewing. The panospheric sensor provides a monocular view, and there was no intent to compare the value to telexplorers between stereo imagery and the immersive mono video used in the research. Also, the video was returned solely for human visual viewing, although machine vision algorithms could be applied to the technology to acquire additional information about the environment.

Approach

Once the decision was made to evaluate immersive imaging applied to teleoperation, it soon became clear that a truly useful method for acquiring immersive views from a mobile robot did not exist. The groundwork of the research was in understanding the flaws of traditional components which was achieved through attempts to design a mobile robot that performed reliable and simple immersive imaging. Eventually panospheric options were considered and viewed as having potential.

To understand awareness and design a perception sensor, the human visual system was reviewed along with those of animals that have "immersive imaging" capabilities. The value of foveal and peripheral vision and the actual angles and resolutions of human vision were compared to the capabilities of available panospheric technology. Additionally, the vision systems of several birds were reviewed to see if lessons from nature might be applied.

Prior studies on awareness and wide field of view imaging for teleoperation and lessons from deployed teleoperated robotic systems were reviewed. Human factors studies on the value of awareness as well as anecdotal evidence of teleoperators provided suggestions on the placement of the sensor on the robot as well as the range of views needed. Similarly, looking at the displays used for teleoperation and simulation environments provided insight into the gaps in immersive imaging and presentation.

Several designs for a sensor were considered that could provide immersive video, and were simulated or implemented for assessment. Graphical simulation was used to model the optic and generate a few images to allow development of display software. Three hardware versions of the sensor were created with lessons from the first creating the second which was refined and miniaturized to make the third (the one used for evaluation). The sensor was designed and implemented for the Atacama Desert Trek which was designed to perform and evaluated telexploration.

The panospheric video was displayed in various formats and on various media. Initially still frames were converted for display on a flat screen. For evaluation of the video during telexploration, the video was enabled for display in an immersive theatre. On the theatre screen, the video was displayed in full theatre views, front views, views with pan/tilt available, raw views, and hybrid views. Telexplorers utilized these views during the operation of the robot, while attempting to successfully traverse unknown terrain. Observations and results were collected and interpreted.

Synopsis of Dissertation

Opting for HyperWide Views

In **Opting for HyperWide Views**, the effect on operator awareness provided by availability of wide-field of view imaging is shown via lessons from the robotic and biological worlds. Understanding the remote environment is necessary for teleoperation, and is achieved in several ways. Although humans have foveal and peripheral vision to aid in achieving awareness, some animals have vision systems explicitly designed for immersive viewing. Early teleoperation work considered wide fields of view, with overall favorable results.

Success in teleoperation has been correlated to the operator's situational awareness, or ability to understand and relate to the robot's environment. Humans use several methods to achieve situational awareness, of which imagery is the most common. To maintain awareness of the remote environment, information from the peripheral regions is commonly supplied to the operator until the environment is held in memory.

Being the de-facto standard in everyday life, human vision is also the standard to which visual data is held by viewers. The human vision system is comprised of two components, foveal and peripheral. Each one serves a different purpose, and has different characteristics. Birds, like many animals, have vision systems designed to provide immersive imagery, which has advantage in avoiding becoming "prey".

Prior studies have been undertaken in evaluating the role of field of view in teleoperation success, but did not utilize immersive video acquisition and display. Studies comparing the relative merits of field of view and resolution were common at the beginning of teleoperation as the imaging technology was not sufficient to offer both, and field of view was considered important for navigation. A few robots have been created for exploration tasks, and the value of immersive viewing is clear from their lessons. Notable examples are the moon-traversing Lunokhod, the Marsokhod experiments in Kilauea, the Dante II expedition in Alaska, and Sojourner Truth on Mars.

Both a sample task and actual telexploration event are used to explain imaging challenges for which panospheric video is shown to be an appropriate solution. In lunar exploration robot design, immersive visualization and panospheric video is a feasible and elegant solution. For the Atacama Desert Trek, panospheric video makes sense, but the differences from a lunar application result in a few design differences. The origins of panospheric imaging and related uses are described, and the format of the panospheric image is explained.

Achieving Full-Surround Video

An imaging solution for **Achieving Full-Surround Video** for exploration is motivated and designed and simulated, with details given of the implementation and the display. Designing immersive imaging solutions present the failings of traditional approaches and motivates an innovative panospheric solution. Creating a panospheric sensor is accomplished through iteration of design and implementation. To use the panospheric image as a normal image, it must be transformed to fit the display media.

The design of the panospheric optic presents dilemmas and options which are addressed here via models, simulation, and iterations of implementation. The raw image from a panospheric camera is awkward for direct viewing by a human operator, but through software transformation, an image can be created that is identical to traditional camera views. Altering the optic requires a corresponding software change, but the display is unchanged and internal changes are transparent to the viewer. Additionally, the image stream can be adjusted to fit the display, whether it be flat-screen monitor, or domed immersive theatre. Ray tracing techniques were used to create sample images in simple domains before constructing sensor prototypes.

A panospheric sensor was implemented and utilized for telexploration with the Atacama Desert Trek. The sensor consisted of the optic, camera, framegrabber, processor, power supply, and software algorithms for regulating acquisition and performing transmission. The sensor was packaged for operation in the harsh environment of the desert, and was reliable and effective throughout the telexploration experience.

During the Atacama Desert Trek, video was displayed in four locations simultaneously, using different methods of display. A 32 seat immersive theatre at Carnegie Science Center in Pittsburgh was the principal display, with all telexplorers observed for this research using only this display. A flat-screen monitor version was used when the color and contrast of the immersive theatre was insufficient to identify potentially dangerous terrain. A big-screen TV projector was utilized at NASA Ames, and video was also passed to an NTSC output for a local cable channel.

Evaluating Panospheric Telepresence

The process of **Evaluating Panospheric Telepresence** involved determining various forms in which to display the video, observing and debriefing the telexplorers, and interpreting the data for common themes and results. A pan/tilt approach was mimicked and juxtaposed with a fully immersive, a raw panospheric, and a virtual-reality style approach for display. Telexplorers performed better with access to peripheral data, but preferred resolution over immersion. Novice operators differed in suggestions and comments from operators familiar with standard teleoperation techniques.

Observations showed that peripheral video and resolution affected teleoperation significantly. The ability to access peripheral views was critical for telexplorers to react to untraversable terrain. The full immersive view was too new and confusing for many operators, who also found the resolution too coarse when covering the full 40 foot screen. Novices and experienced operators had different expectations, but not significantly different capabilities. Understanding the locomotion system was the main technical advantage of experienced operators.

The thesis assertion that video conveying the entire surroundings of a teleoperated mobile robot enhances situational awareness needed to react to untraversable terrain while exploring was evaluated and from the results of the observations deemed to be valid.

Retrospections

The accomplishments of the research and arenas for future work are listed, and **Retrospections** on the contributions lead to insights from the author.

This research accomplishes the innovation of panospheric video for remote operation of mobile robots. A panospheric sensor was developed and demonstrated on a robot performing telexploration. Immersive imaging as a means of achieving situational awareness was evaluated in varied display configurations.

Future work exists in upgrading the sensor, advancing the single-image-multiple-viewer capability, porting traditional machine vision algorithms to the immersive domain, and improving immersive displays. The sensor could benefit from continued advances in imaging technology and changes suggested during deployment in the desert. Enabling multiple viewers, even at distant locations from each other, to interact with the video as a team would further increase the utility of panospheric imaging for telexploration. With feature detection, algorithms such as range from motion flow or image sequence, obstacle avoidance, integrated positioning can be adapted from the machine vision library. Upgrading domed displays to directly accept panospheric imaging formats or developing immersive personal displays is essential to the continued growth of immersive imaging in robotics.

Along with adding a previously absent technology to the arsenal of the robotics designer, the contributions of the research stem from the ability of the robot and/or its operator to utilize panospheric technology to simultaneously view in all directions. Telexplorers do not need to plan a motion or pan/tilt sequence in order to achieve the desired views, and while passing an object the front, side, and back of the object are automatically imaged. Each operator can customize the display of imagery to their liking, rather than having to adapt to the imagery of the robot.

Utilization of panospheric video for robotic telexploration is an appropriate solution, even though current technologies fall short of achieving the incarnation of human vision. In its initial deployment, panospheric technology was shown to have significant value over traditional methods of image acquisition. Solutions to any problem previously discarded due to complexity should be reexamined in light of technological advances. Even when immersive imagery is made equivalent to human vision, acquiring the imagery is only one part of replicating human visual-based perception.

Chapter 2

Opting for HyperWide Views

In **Opting for HyperWide Views**, the effect on operator awareness provided by availability of wide-field of view imaging is shown via lessons from the robotic and biological worlds. The imaging challenges for telexploration are explained using a sample task and an actual event. An innovative imaging solution, the panospheric optic, is explained both in terms of its origins and its format.

Success in teleoperation has been correlated to the operator's situational awareness. Understanding the environment about the robot allows the operator to work more effectively and safer. Imagery is a crucial tool to understanding the environment. Prior studies have been undertaken in evaluating the role of field of view in teleoperation success, but did not utilize immersive video acquisition and display.

As humans are the principal operators of remote robots, human vision is the standard to which visual data is held. When it comes to achieving situational awareness, however, other animals have vision systems that are more ideal. Combining the human standard with immersive awareness is desired, but beyond the state of current technology.

Both a sample and actual telexploration task present visualization challenges for which panospheric video and immersive displays are shown to be an appropriate solution. The format of the panospheric image is explained.

Situational Awareness in Teleoperation

Situational awareness describes the state of having knowledge about one's environment. In everyday life, humans use a variety of ways to acquire and maintain information about the environment. During teleoperation, the ability of the user to maintain situational awareness is paramount to successful completion of the task, and the dangers of losing awareness of the robot's environment are well documented.

Humans predominantly use visual cues for achieving awareness, and imagery provides crucial contributions to situational awareness in robotics as well.

Dangers of Losing Awareness

Robot operations have been thwarted during teleoperation when situational awareness is lost by the operator. The problems have included vehicle accidents that ended operations to considerable loss of operation time spent trying to understand a particular situation and return to planned operations.

In the late 1980's, tests done at Sandia National Laboratories had operators in control of off-road vehicles. The operator was tasked to drive over the desert sand dunes either to get from point to point, or to perform general reconnaissance. The imagery provided to the operator came from a 40degree field of view camera fixed to the body of the robot. From the imagery alone, operators were unable to determine the slope currently being traversed by the vehicle, as the dunes were large enough that the local horizon always looked safely level as seen from the camera mounted on the vehicle. The result was numerous occasions when the vehicle was rolled over [McGovern87].

In 1994, NASA Ames Research Center's Intelligent Mechanisms Group performed tests with the Marsokhod vehicle on the lava fields of Kilauea in Hawaii. The vehicle was fitted with a conventional camera fixed to the body with a forward field of view. In addition, a camera was attached to the wrist of a manipulator arm. The intent of the operations was to perform remote science by locating and investigating interesting geological features. The majority of the features of interest were ravines or canyons in the lava where strata were visible. Frequently the robot was driven into a narrow canyon and needed to be turned around to be driven out. From the imagery on the front of the robot, it was not possible to see what was behind the robot, and on numerous occasions, the support team in the field had to take over control of the robot and perform operations to extricate from close quarters [Fong95].

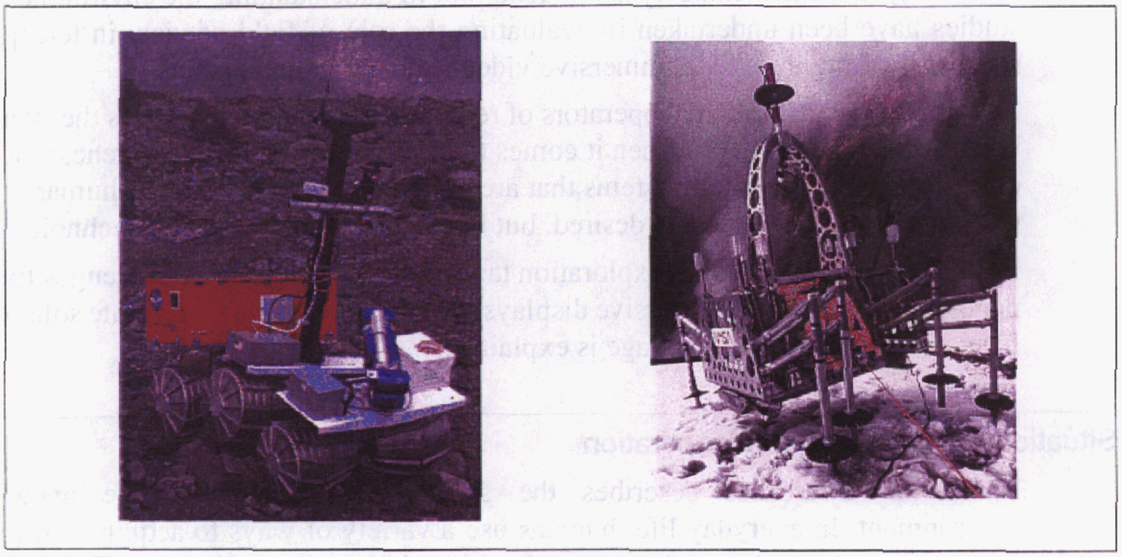


Figure 2-1: The Marsokhod robot as deployed during the Kilauea field experiments, and Dante II descending into the crater of Alaska's Mount Spurr

Also in 1994, the Dante II robot of Carnegie Mellon University's Robotics Institute was teleoperated into the crater of the active Mount Spurr volcano in Alaska. Equipped with a 2D scanning laser, and multiple cameras positioned over the structure, the robot had a fairly good capability for providing awareness. However, not all the views were available simultaneously, and switching from view to view for nearly every footstep of the walking robot was taxing on the operators. The descent was accomplished successfully, and an attempt was made to climb back up the slope. At some point, the operators were unable to detect that the ground had shifted under the robot's weight and subsequent commands resulted in tipping the robot onto its side [Krotkov94].

Ways to Convey Situational Awareness

In everyday life, humans use situational awareness constantly to get from place to place without experiencing difficulties. Most of the time, no conscious effort is spent on interpreting the environment and making decisions. As an example, students routinely travel down crowded hallways while engrossed in conversations with companions, and rarely have collisions with other students or walk into the wrong room. Some of this success is based on memory of having been the way before, and some of it is based on keeping track of the dynamic conditions around them and making appropriate command decisions. Attempting to perform comparable tasks with a teleoperated robot is astoundingly difficult and tedious, both because the perception of the remote environment is not as easily obtained, and the mechanisms are frequently not as agile in the event of failed perception. Slowing down the speed of operations alleviates the agility requirement, but the requirements in perceiving the environment and the robot's place within it remain.

Significant amount of research in robotics has gone into mapping and position estimation techniques. With the map and the robot position, the teleoperator is able to make plans about what to expect from the terrain, and determine what commands to send to the robot. Navigation based on map and position is used by humans in instances like the "you are here" arrow found on shopping mall directories.

When following a route, robots and operators can share locations of landmarks to verify progress. Just as most sets of directions to a new location contain phrases such as "turn right at the third stoplight", recognizing landmarks provides a way for teleoperators to operate without a definite map or global position. In this case, the operator must rely on perception sensing from the robot to react to the terrain, but can make high-level command decisions based on the landmarks.

Contributions of Imagery to Situational Awareness in Robotics

Robots ignore most of the senses available to humans, including hearing, smell, taste, and touch in favor of various forms of "sight" and kinetic knowledge. Although there are robots that use acoustic information to understand what materials are comprised of (via tapping the surface) or for localizing where sound is originating, the general lack of auditory information means that robots cannot tell that something is coming around the corner. Many forms of "sight" are employed, ranging from visual perception, laser scanners, sonar rings, and radar arrays. These sensors provide the robot and operators with a description of the features in the environment at a particular location at a particular time [Martin95].

Using kinetic knowledge such as odometry, inertial measurements of accelerations, inclinometers, and force sensing allow robots and operators to augment the sight based

10.300	9.099	11.859	3.796	8.300	8.951	12.115	3.534
1.000	8.874	12.402	3.272	1.000	8.874	12.698	3.011
5.600	8.951	12.985	2.749	12.000	9.099	13.241	2.487
14.700	9.309	13.451	2.225	5.800	9.565	13.599	1.963
8.800	9.852	13.676	1.702	10.300	10.148	13.676	1.440
22.400	10.435	13.599	1.178	10.500	10.691	13.451	0.916
1.600	10.901	13.241	0.654	1.500	11.049	12.985	0.393
1.400	11.126	12.698	0.131	1.400	11.126	12.402	6.152
1.500	11.049	12.115	5.890	1.500	10.901	11.859	5.629
8.100	10.691	11.649	5.367	6.700	10.435	11.501	5.105
6.600	10.148	11.424	4.843	6.400	9.852	12.424	-1.702
3.500	9.565	12.501	4.320	9.300	9.309	12.649	4.058
9.600	9.099	12.859	3.796	5.700	8.951	13.115	3.534
1.000	8.874	13.402	3.272	1.000	8.874	13.698	3.011
6.500	8.951	13.985	2.74	1.400	9.099	14.241	2.487

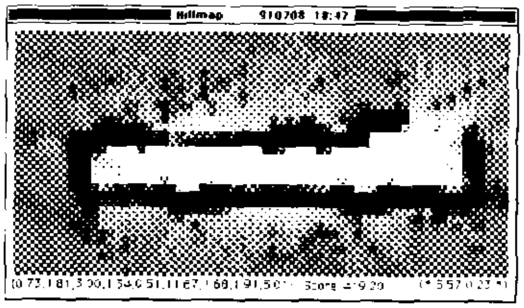


Figure 2-2: Raw data excerpt of sonar array output and visual display of sonar data (courtesy Hans Moravec)

perception and gain more information about the environment. However these kinetic sensors provide information only about the current position of the robot, not about what the robot is about to experience. With only kinetic knowledge, the robot is being operated blindly.

Robots performing exploration operate in a domain unknown to the operator, and presumably to the robot. As a result, it is not possible to memorize the world and operate blindly. Robots that follow lines in the center of hospital hallways, or that take repetitive paths across a factory floor can afford to operate blindly with only a minimal sight range for preventing collisions. The imagery obtained by the robot is transmitted back to the operator who must interpret it and make decisions. If the robot is acquiring sonar, laser, or radar data, the data must be transformed into a format more interpretable by the human teleoperator than reflectance values. This is commonly some form of visual presentation as shown in Figure 2-2.

In the case of remote geology, a particularly important telescience goal of NASA planetary robotics, visual imagery is the basis for success. The general operating pattern is to select an area of probable interest from satellite or aerial photography, travel to that area, acquire a panorama to determine position in the region and make selections of areas to visit, and then travel to the specified regions and perform the science experiments. Without the overhead imagery to select interesting regions, the ability of the robot to acquire the panorama is significantly more important. This is the case when exploring planetary bodies such as mars, outer moons, or asteroids where satellite imagery is not likely to be available.

Studies of Imagery and Teleoperative Awareness

Early teleoperation research studied the value of imagery and fields of view on the operator's effectiveness and comfort. Anecdotal experience from more recent robot systems confirms that lack of imagery correlates directly with lack of awareness and lack of success. These modern systems also show the desire for immersive imaging, and that traditional solutions are not ideal.

Early Studies of FOV and Teleoperation

One of the earliest studies on the effect of FOV to teleoperation was a military feasibility study [Eveleth76], which noted that operators of a remotely controlled vehicle were unable to detect target arrays even at close distances due to the narrow

FOV. A decade later, tests at Fort Knox in the Advanced Ground Vehicle Technology (AGVT) Concept Evaluation Program suggested that wide FOVs were needed to maintain spatial orientation with respect to terrain landmarks, especially in unknown terrain. This work supported an unrelated study which found that for the specialized case of roadway driving, on-board operators were safe at 25km/hr with only a 4degree FOV [Gordon66]. However, in driving off-road, without the center-line available for reference, operation with this visual restriction is impossibly tedious.

One of the first attempts at panoramic displays was also done by AGVT, in which three 60degree cameras captured the front hemisphere and were displayed on 45" rear projection televisions in a semi-circle about the operator. In this work, it was found that peripheral information is extremely useful for close-quarter maneuvering and cross-country driving [Almaula86]. Similar results were obtained at Sandia National Laboratories from operators attempting to maneuver a Jeep Cherokee in a parking lot using both 40^0 and 120^0 views. The control station for this work is displayed in Figure 3-13 [McGovern80].

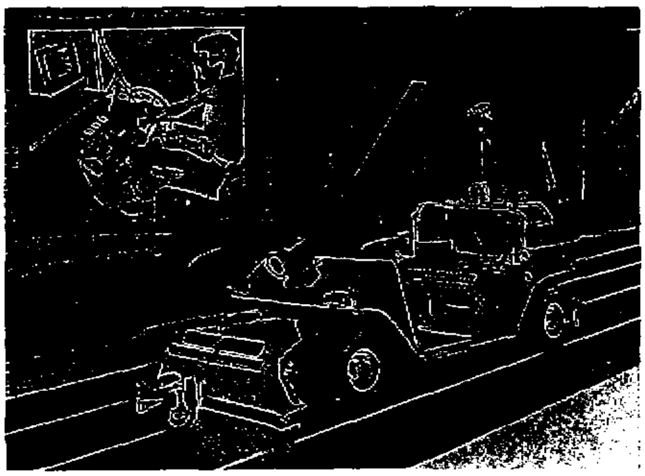


Figure 2-3: Experimental setup at Army Research Laboratory for testing FOV and teleoperation

In tests by the Army Research Laboratory (ARL) 29^0, 55^0, and 94^0 FOVs were compared to hands-on operation. Due to the distortion of imagery acquired through the widest lens, performance of remote driving was most successful for the middle case, except for obstacle avoidance in which the wide angle was superior. The narrow FOV was always inferior to the other two [Glumm92]. Wide FOV was initially not pursued in teleoperation because it was suspected to cause the motion sickness experienced by 10 to 60% of operators, depending upon the type of control station. However, it was shown that the symptoms of motion sickness can occur with a FOV as low as 7.5^0, and motion sickness is now attributed to disparity in the visual and vestibular senses (those responsible for maintaining balance and orientation to vertical) [Alexander90].

Lessons from Modern Robots

Robots developed and deployed in the 1990's have the advantage of better computer technology, and were able to process more complex algorithms, utilize higher bandwidth transmissions, and provide the operator with more data. As a result, the robots were able to move faster in the environments and were even able to begin to aid the operator in interpreting sensor data and perform basic safeguarding. The Scamp robot of University of Maryland's Space Systems Laboratory, and the Ratler robot and Unmanned Ground Vehicle (UGV) robots at Carnegie Mellon University's Robotics Institute provided further lessons in awareness and visualization for teleoperated robotics.

Scamp is a spherically shaped submersible camera platform and was the precursor of NASA's recently deployed in-orbit AERCAM robot. Ratler is a four wheel skid-steer robot based on a rocker-bogey platform with a stabilized central mast, and was the precursor to the Nomad robot used on the Atacama Desert Trek. Both these robots utilized a fixed forward camera with a 60degree field of view. In independent teleoperation experiments using Scamp and Ratler, a similar task was required of the operators: to navigate through an obstacle course. In the case of Scamp, the obstacles were hula-hoops suspended at various depths that the operator was supposed to pass through. In the case of Ratler, the obstacles were tires lying on the ground that the operator was supposed to avoid.

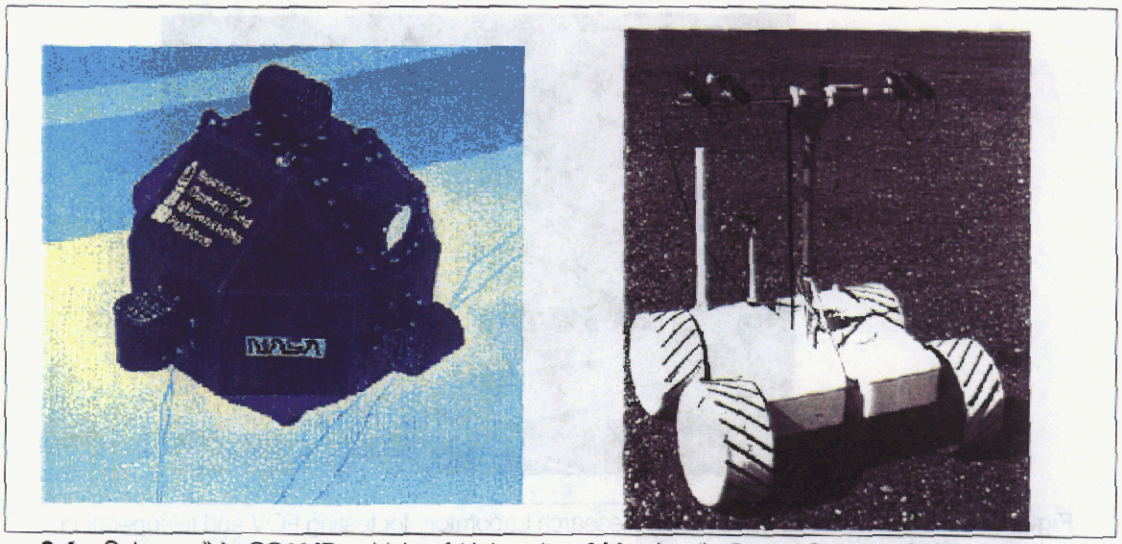


Figure 2-4: Submersible SCAMP vehicle of University of Maryland's Space Systems Laboratory and the RATLER vehicle developed at Sandia National Laboratory and augmented at CMU's Robotics Institute

The principal problem experienced by the operators was that when the robot got near to obstacles, they were no longer in the field of view of the camera. In the case of Scamp, operators would get the camera through the hula-hoop, but hit it with the body of the robot, and occasionally get hung up for several minutes trying to figure out why the robot would not go forward. With Ratler, the operator would often catch a tire on one of the back wheels, and experience the same frustration at not being able to see the problem. These experiences continue to confirm that immersive imaging of some form is required for teleoperation success [Hunt96][Krotkov94].

The vehicles of the UGV program were off-road military ambulances (HMMV's) that were fitted with stereo cameras. The primary operating mode of the UGV's was a semi-autonomous mode, where the operator specified the destination way points and the robot undertook the responsibility to get from point to point. The principal perception method was multicamera stereo imaging which allow the robot to compute terrain maps and make decisions about local traversal routes.

The cameras were mounted on a pan/tilt mast. When the vehicle turns, the camera head is panned in the direction of the turn. When the vehicle crests hills or comes to the bottom of valleys, the camera head is tilted so that images are obtained of terrain that is ahead of the vehicle. The desired field of view and number and viewing angle of cameras was designed for machine vision processing, but indicates the need for operators (human or robot) to see in extended viewing directions during operations. On a side note, in a recent demo one of the servo motors in the pan/tilt head failed, punctuating the reliability concerns about moving parts [Brumitt97].

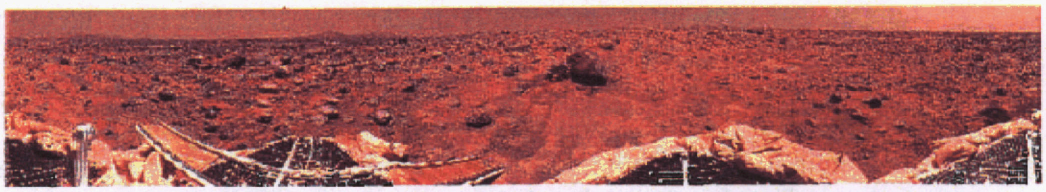


Figure 2-5: Image of mars created by a sequence of images from the Mars Pathfinder pan/tilt camera.

The Sojourner rover which landed on Mars in 1997, was teleoperated through the use of panoramic imagery and stereo imagery acquired from the Mars Pathfinder lander. A pan/tilt camera was elevated upon a mast and acquired a multitude of images around the entire azimuth. These images were then patched together to form a panorama. However, due to the parallax issues of the camera and the changing position of the center of projection from the pan/tilt motion, it was not possible to seam the imagery together without introducing warping. The complete panorama is shown in Figure 2-5 and a seam in the imagery (the bottom left corner of the panorama) is shown in Figure 2-6.

Figure 2-6: Visible seam in the Pathfinder panorama.

Teleoperation and Biological Vision

For telexploration, both telepresence and awareness are required. Human vision is the de-facto standard for telepresence, and the characteristics of human vision are described. With vision more geared to forward viewing, humans lack in awareness. Many biological creatures have vision systems more devoted to immersive awareness, and the vision systems of several birds are detailed. When performing teleoperation, some traits of vision are more important than others. Technology for achieving telepresence and awareness during teleoperation is not yet sufficient to replicate immersive imaging at the human standard.

Statistics on Human Vision

The human visual system is significantly more impressive than even the most advanced of cameras. Inanimate means for achieving the resolution of the foveal region, the ability to rapidly look from point to point, the total field of view of the eyes, the dynamic range of the retina, and the exposure control of the pupil can (in some cases) be achieved individually, but not in combination. The human eye has two field of view components, foveal and peripheral. It also merges together images from both eyes to produce a single apparent display. Binocular vision is used to obtain depth information of the environment.

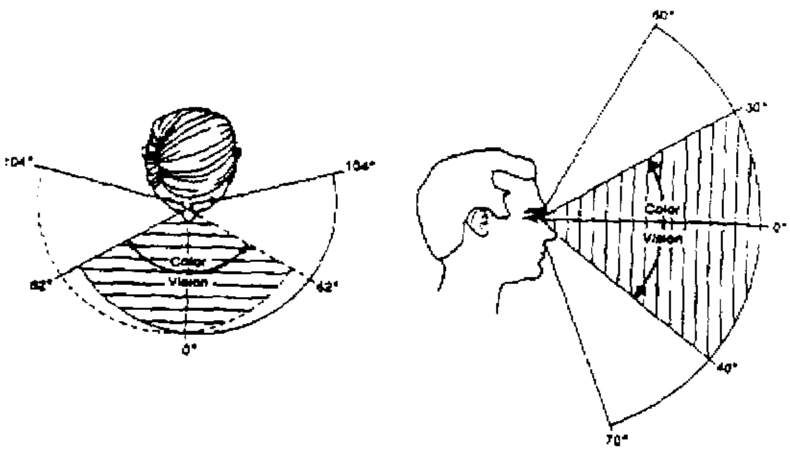


Figure 2-7: Field of view ranges of human vision showing color vision ranges

The field of view of the human eyes are shown in Figure 2-7. The average human field of view without allowing eye or head motion is 204^{0} in azimuth and 130^{0} in elevation (slightly oriented downward), and 140degrees of the azimuth are overlapping, or binocular vision. Of the entire range, only the center 124^{0} by 70^{0} is in color and the rest of the peripheral view is black and white [Kress88].

The resolution of the foveal region at 20/20 vision is sufficient to resolve a quarter size object at a distance of 3km. This assumes optimal contrast between the target and the background and no atmospheric effects. With the human foveal resolution taking up only 7degrees of the field of view, camera systems that attempt to acquire identical resolutions over a larger range have a significant challenge. In 1988, Gary Kress and

Haren Almaula documented the properties of both human vision and standard teleoperation visualization. The trade-off between FOV and resolution was illustrated by showing that a camera yielding normal human acuity (20/20 vision) with a 7^{0} view degrades to 20/150 when used to capture 90^{0} [Kress88]

Visual Systems of Birds

Human vision is normally thought to be highly advanced, but it is advanced only for the environment in which we operate. As a creature that can be a predator and one that works with tools significantly, the heavy fraction of viewing angle for binocular field of view is extremely useful. Yet as humans can also be prey, the wide peripheral field of view enables awareness of some degree to both sides. Combined with the angles the neck can turn to the sides and upwards, it is possible to quickly scan the entire surroundings by turning the head.

Birds such as owls and eagles tend to have vision that more closely resembles human fields of view, with the snowy owl having roughly a 200degree azimuth with 100degrees of binocular vision. The owl however can turn its head over 180degrees in either direction enabling it to have binocular vision in all directions rather than just peripheral vision in the back hemisphere. The golden eagle has a larger monocular vision range (about 310degrees) but a much smaller binocular range (50degrees). However, its foveal resolution is estimated to have 10 times the acuity of humans, enabling them to detect breathing in a field mouse from an elevation of 2km [Freethy82].

The vision of non-hunting birds contrasts to birds of prey. The common pigeon for instance has a 340degree azimuth of vision, but only 20degrees of binocular vision. In addition, their eye has two foveal regions. One that is in the forward direction for use with stereo, and one that is to the side and slightly downward. The apparent bobbing and nodding of the head is the pigeons way of moving this foveal region about to track objects that might be sneaking up on it, while still keeping its prey in focus as well.

The woodcock, and a few other birds that commonly spend time with their head pointed downward digging in marshes for food have a rather astounding field of view. Each eye has a 200degree field of view, giving it a full 360degree of azimuthal coverage with both the forward -and backward- 10 degrees in a binocular region. This allows the bird to keep track of where to place its bill, as well as to keep track of predators that might be sneaking up on it, especially from above or behind.

Vision Traits that Affect Teleoperation

To efficiently teleoperate a remote robot, some characteristics of vision are more important than others. As shown in the studies and examples, field of view, whether via the ability to pan/tilt to the desired location or from immersive fields of view, is extremely important in understanding a robot's surroundings. The ability to see more field of view allows for quicker understanding of the world, increasing the speed of decisions. In the same way, the update rate of the imagery increases the understanding of the current environment, and increases decisions and overall operation speed.

Resolution provides for the ability to distinguish the line between danger and safety, and allows for precision navigation in cluttered or close environments. The trade-off of resolution and FOV on operator visual interfaces for remote operations was evaluated

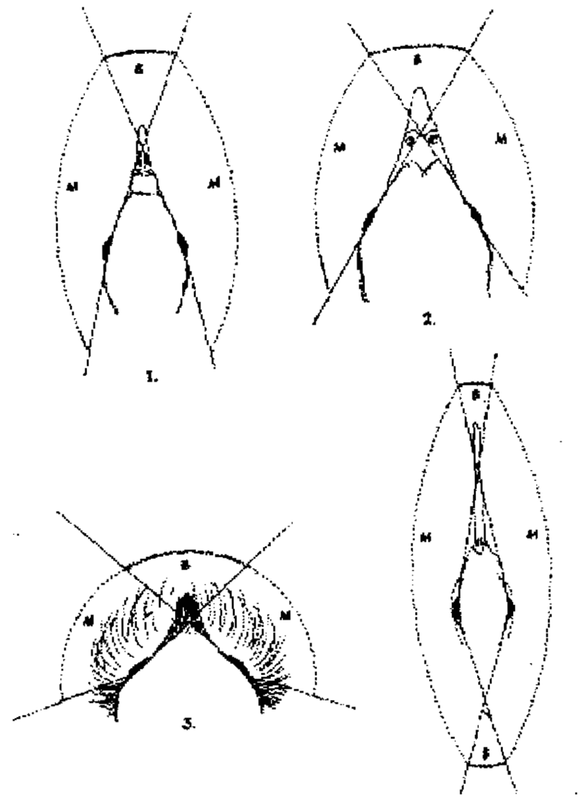


Figure 2-8: Monocular and binocular fields of view of four birds: 1. pigeon, 2. golden eagle, 3. snowy owl, and 4. woodcock (yes the woodcock does have binocular vision directly behind its head!).

in 1986 at Sandia National Laboratories and showed that resolution was important for identification of objects, but that FOV was important for locating objects [Miller86]. Remote science operations have suggested that high-resolution imagery comparable with human foveal vision substantially aids in understanding.

Color and dynamic range make the scene more understandable, increasing confidence. Being able to see into shadowed areas, or being able to distinguish the grey sidewalk from the green grass enables operators to evaluate the safety of a proposed route, and enabling short cuts to be taken that increase operation efficiency.

Technology Compared to Biological Vision

Telepresence, both in robot operations and entertainment is increasingly held to human vision standards. While television is the most common display source for visual data, it is not at all immersive, and has rather poor resolution (640 x 480 is commonly used when recording digital media onto VHS tapes), especially when broadcast on a projection television or on a 72inch screen. In contrast, the recent purely digitally animated movie, Toy Story, used a resolution of 1526 x 922 pixels per rendered frame to look appropriate in movie theatres, with each pixel taking up 1/4 of an inch on the movie screen [Disney96]. While the movie screens do have higher resolution, the proscenium style presentation allows for no display of peripheral imagery as part of the experience.

Virtual reality headsets would seem to have the ability to offer both resolution and immersion, but the technology at this time is really not much more immersive than television, with only slightly better resolutions. Perhaps the closest acquisition and display system to mimicking human vision is the OmniMax film standard which uses three 70mm movie projectors to seamlessly fill a curved screen that covers 140degrees of horizontal by 70degrees vertical (about the range of human color vision).

Digital cameras lack the resolution and framerate to compare to the human eye over an extended field of view, but are getting closer to duplicating the dynamic range and color sensitivity. Film based cameras can nearly duplicate the human vision system, but the developing process keeps the data from being real-time, and the media is infeasible for teleoperated robots.

Imaging Challenges for Telexploration

To understand the difficulties in achieving visualization for robotic telexploration, imaging challenges for both a sample task and actual event are detailed. The LunaQuest is a long distance and duration traverse across the lunar surface visiting historical sites and interesting geological features for the purposes of edutainment. The Atacama Desert Trek was a demonstration of technologies for lunar exploration, with the stated goals of 200km of teleoperated distance across a planet-like terrain, operations predominantly by novices, and performance of telescience experiments. Nomad was the robot used during the Atacama Desert Trek, and its key technologies and specifications are listed.

Figure 2-9: A lunar surface exploration mission for education and entertainment (courtesy LunaCorp)

The LunaQuest

Exploring the Moon offers entertainment potential, as well as scientific and educational interests. The LunaQuest is a two-year, 1000km excursion to visit historical lunar sights in and around the Sea of Tranquility. The entertainment is provided in the form of live telepresence to theme park enthusiasts, television documentaries, and collections of image libraries for CD-ROM or Virtual-Reality exhibitions. The primary technical challenges in the rover design are providing continuous and rich live video, providing robust mobility on diverse terrain, and operating for two years in the inhospitable lunar environment. While a single rover could accomplish a traverse between all these sites, a pair could make the trek more entertaining and reliable by providing off-board views of one rover from the other.

The LunaQuest begins in the vicinity of Tranquility Base, the site of the first manned landing by Apollo 11 in 1969. In short proximity are the impact crater of Ranger 5 and the landing site of Surveyor 8. About 500km away in the highlands at the north of the Sea of Tranquility, lies the next destination, Apollo 17. Accompanying the astronauts on this mission was a version of the Lunar Rover Vehicle. The last stop, and the true quest, is the Soviet Lunokhod 2. This "ancestor" of the LunaQuest rover was the last vehicle to be teleoperated on the Moon, and traversed about 40kilometers before becoming mired in a crater. Its exact location is unknown [Katragadda96].

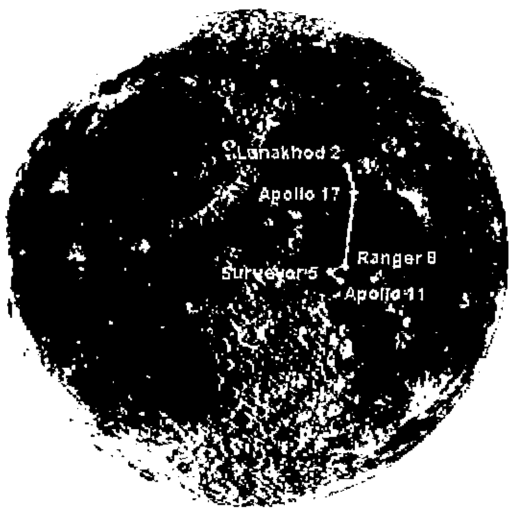


Figure 2-10: The trek route for the LunaQuest historical site tour of the lunar surface

The lunar environment is difficult for several reasons, but has many advantages for robot operations as well. The temperatures on the lunar surface range from that of superheated steam to liquid nitrogen over the day/night cycle. The radiation is low, but must be survived over a two year time period. The lunar dust which comprises the surface is extremely fine, a nearly ideal insulator, and has electro-static creep along conductive surfaces. More favorably, the lunar surface is devoid of vegetation and moving obstacles which thwart current perception and navigation techniques. A direct line of sight between the lunar near-side and the Earth is always available for maintaining communication. Without a filtering atmosphere, winds, or clouds, sunlight is a rich and guaranteed power source. The time delay between the Earth and the Moon is small enough (about two seconds round trip) for teleoperation of a lunar vehicle to be plausible.

The launch/landing scenario calls for two rovers to be launched via a Proton rocket and delivered to the lunar surface on a Phobos class lander. The rovers are stowed on the lander during launch in a vertical position with the wheels attached to ramps that lower to the lunar surface for deployment. For stability of the lander during descent, each rover must weigh less than 250kg and fit within a 2m x 2m x 1.5m envelope for travel.

The need for low power, low mass and high reliability are all tied together [Whittaker9X].

Challenges for Imaging

The primary product of the LunaQuest scenario is delivered imagery from which a lunar telepresence is created. In addition, imagery is used for achieving awareness by the rover operators. In general, the desired targets to image are the surrounding terrain, the other rover, the Earth and stars overhead, and interesting geological features of the landscape.

The goals for the imaging system generate challenges:

- for planning and safeguarding, a wide field of view is needed to the front
- for diagnosis or extrication from hazards, the terrain around the wheels and back of the rover should be visible
- for observing artifacts the rover is passing or circling, side views are needed
- for obtaining position from the stars or viewing earth, overhead views are needed
- for providing edutainment, the imagery needs color, video speeds, and display in real-time
- for telescience discrimination of geological features, high resolution is needed

To meet all challenges successfully, the robot must capture imagery to the front, sides, back, near-ground, and overhead, or more succinctly, immersive imagery.

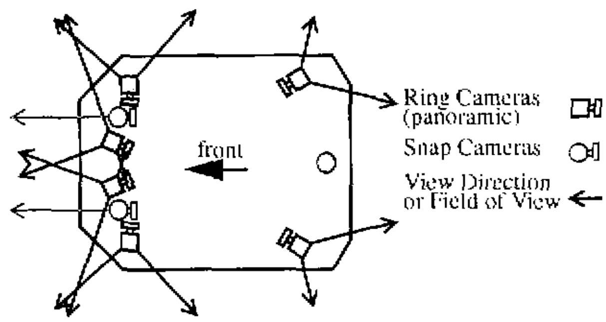


Figure 2-11: Configuration for the LunaQuest imaging needs using an array of cameras

The initial camera configuration, shown in Figure 2-11, used nine cameras to acquire the panorama, forward views, and overhead view. This layout had no rotating mirrors, pan-tilt gimbals, or zoom lenses. Only two camera types were utilized, each with the same optic (a 90x60degree field of view). Six of the cameras were used to create a panorama from -45 to +15degrees, and the other three cameras were used to acquire higher resolution snapshots with color filter wheels. This configuration required too much power, too much mass, and added too much complexity in terms of frame grabbing, camera calibration, and cabling. It also failed to provide an adequate vertical range of imagery to enable near-ground views for extrication from hazards.

The Atacama Desert Trek

The Atacama Desert Trek was funded by NASA Telerobotics to demonstrate robotic technologies relevant to planetary exploration. The objective was to develop and demonstrate a robot capable of long distance and long duration operations. The goal of the trek was to travel 200km under both teleoperated and semi-autonomous intercontinental control. A secondary goal was the performance and evaluation of remote science. The trek occurred in the Summer of 1997 from June 15th to July 31st, with the week of June 18th to June 25th devoted to telescience [Whittaker98].

Located in northern Chile, the Atacama Desert offers terrain analogous to planetary conditions. The landscape includes craters, rocks, and loose sand. There is no vegetation due to the lack of precipitation and the high mineral content of the soil. The terrain is rugged and includes obstacles that will be impassable for the robot. The Atacama Desert is located in the same time zone as Pittsburgh, simplifying co-ordination of communications and allowing operations during regular daytime hours.

The Atacama Desert Trek involved public participation through operations at Pittsburgh's Carnegie Science Center, and also at NASA Ames Research Center. Nomad's safeguarded teleoperation, combined with a rich, interactive user interface was designed to allow novice operators the opportunity to operate Nomad safely.

Figure 2-12: Nomad in the Atacama Desert

Nomad

The robot developed for the exploration was named Nomad. The charter for Nomad was to be highly reliable and robust, and operate as though on a remote planetary surface. As a testbed for planetary relevant technologies, Nomad had innovations in locomotion, imaging, communication, and safeguarding. As a planetary exploration analog, Nomad was intended to fit aboard a lunar lander inside the nose-cone of a rocket and as a result had design constraints on mass, dimension, and center of gravity.

The locomotion system was unique in its ability to transform in size once out of the constraining nose-cone, increasing the stability and propulsion. The chassis expands, compacts, and steers by driving two pairs of four-bar mechanisms, one on either side of the robot. This capability not only allowed Nomad to adapt its footprint to terrain, but also enables steering by differential actuation of the two deployment motors. The "deployed mode" improves Nomad's stability and propulsion over terrain, and averages the excursions of the body for smooth motion and consistent operation of on-board sensors and computing.

Maintaining a high-bandwidth communication link between Nomad and a satellite (or to Earth from the lunar surface) involved continual precise pointing. Using omnidirectional antennas restricts the bandwidth and range due to the limited on-board power, and so Nomad used an actively pointed high-gain antenna to achieve high data rate communication over extended ranges. The antenna is steered on a mechanism capable of reacting to the slew rates created by the vehicle motion. The pointing vector was obtained using a combination of inertial measurements and positioning data.

To be reliable while operated by novices, Nomad was designed to be semi-autonomous, with significant safeguarding capabilities. Two pairs of wide angle stereo cameras were used to generate range data and reconstruct terrain for the region in front of Nomad. A line scanning laser was mounted on-board to serve as an emergency stop for ditches that the stereo might fail to detect. Inclinometers allowed Nomad to halt its progress if the slopes of the terrain were approaching vehicle stability limits. The safeguarding system would decide if the direction commanded by the operator was dangerous to the vehicle and adjust the speed or turning angle of Nomad appropriately to avoid the detected hazard [Whittaker9X].

Challenges for Imaging

The telepresence experience of the Atacama Desert Trek based on the acquired imagery was a primary product of the project. In addition, imagery was used for achieving awareness by the rover operators. In general, the desired targets to image are the surrounding terrain, and interesting geological features of the landscape. For telescience, operators and scientists need to be able to select landmarks, and navigate the robot from spot to spot quickly and efficiently

The goals for the imaging system generate challenges:

- for planning and safeguarding, a wide field of view is needed to the front
- for diagnosis or extrication from hazards, the terrain around the wheels and back of the rover should be visible
- for tracking landmarks and maintaining global plans, views of the skyline are needed
- for enhancing the experience of the audience, those watching the telexploration need to have a way to interact and views to the sides of the robot are desired
- for providing edutainment, the imagery needs color, video speeds, and display in real-time
- for telescience discrimination of geological features, high resolution is needed

To meet all challenges successfully, the robot must capture imagery to the front, sides, back, and near-ground. Furthermore, the imaging hardware needs to be rugged and reliable to survive the dusty and windy conditions of the Atacama. Overhead views are not important, as long as the imagery extends enough above the horizon to enable utilization of peaks as navigation landmarks.

Panoramic imagery with an elevation range from -70degrees to +30degrees with respect to the horizon is suggested. The +30degrees is derived from the expected slopes the rover will encounter during the traverse (20degrees) and the desire to see skyline features (about 10degrees). The near-ground angle was selected to ensure that the camera would be able to image both the front wheels given the expected placement of the sensor to one side on the mast. In the actual deployment, the sensor was moved to a central location, and -60degrees would have achieved the desired properties.

The Panospheric Alternative

To meet the visualization challenges of telexploration, traditional solutions will not suffice. However, an innovative sensor design, the panospheric, is capable and appropriate. While never utilized for remote teleoperation, the panospheric concept has existed for tasks such as surveillance, and multiple implementations for achieving panospheric imaging have been developed. As the human vision system is not accustomed to hyper-wide fields of view, the imagery of the panospheric sensor is difficult to interpret, and the format and the features of the raw panospheric image format are described.

Traditional Solutions and Their Shortcomings

To achieve immersive image acquisition in a manner that is reliable and has a low design cost, none of the traditional imaging solutions are acceptable. Pan/Tilt heads cannot instantaneously achieve immersive views. Camera arrays require registration and impose substantial cost on robot complexity. Fixed cameras that require moving the robot to obtain a desired view degrade efficiency and decrease safety.

In the majority of applications that require imaging in varied directions, a pan/tilt camera system is employed. While this enables acquisition of all the desired views, it is usually not possible to perform a continuous slow pan to the left, as eventually the limit of the cables are reached and the camera must swing 360degrees to the right to continue the pan. This singularity is undesirable if there are dynamic objects in the remote environment.

The pan/tilt system is also inadequate because it cannot look in two directions simultaneously. If there is a desire to both view to the front of the robot for operation purposes, and to the side of the robot to gather information about a feature being passed, the pan/tilt mechanism must be pivoted back and forth. It is also possible to completely miss features of the environment because the camera is never commanded to point in the appropriate direction. If constant panning is performed to acquire immersive views as often as possible, the reliability of the mechanism also begins to come into play on long duration deployments, or for deployment in difficult environments.

Similar to panning the camera, some systems have elected to image mirrors that spin around. These circular scanning cameras are better at acquiring simultaneous immersive imagery, but the result of imaging while the vehicle is moving leads to a slight helix pattern, and if the terrain is rough, the image sequence becomes harder to register together into a smooth sequence of immersive images. Again, the reliability of a constantly spinning mechanism also becomes an issue over long cycles.

Replacing the motion with an array of cameras allows for true simultaneous acquisition of immersive imagery without the reliability issues of moving parts. However, the addition of numerous cameras increases the system complexity. There are more sensors (cameras), more cables, more connectors, more requirements for frame-grabbers, and even issues with mux-ing the imagery together for transmission. The overall cost on robot complexity may be prohibited by the design requirements to limit power or mass or volume of the robot. In addition, in order to display the imagery, some registration of the imagery and calibration of exposure and gains and orientation of the cameras must be employed, increasing assembly complexity.

A fixed camera is light and reliable, but cannot provide all the views without considerable vehicle maneuvering, which may not be feasible. Fixed cameras that do not acquire wide fields of view are not exceptionally useful for achieving situational awareness or for providing telepresence, but many robots are forced to use only a single camera by the constraints of the design, and do not have the volume for a pan/tilt mechanism to maneuver.

The Panospheric Innovation

The ideal imaging solution for reliable and low cost robotic telepresence and awareness is a single fixed camera that acquires all the views. While this sounds implausible, the innovative panospheric sensor provides such a solution. Through the use of reflective and refractive optics, a full panorama with extended elevation angles can be imaged from a single camera location.

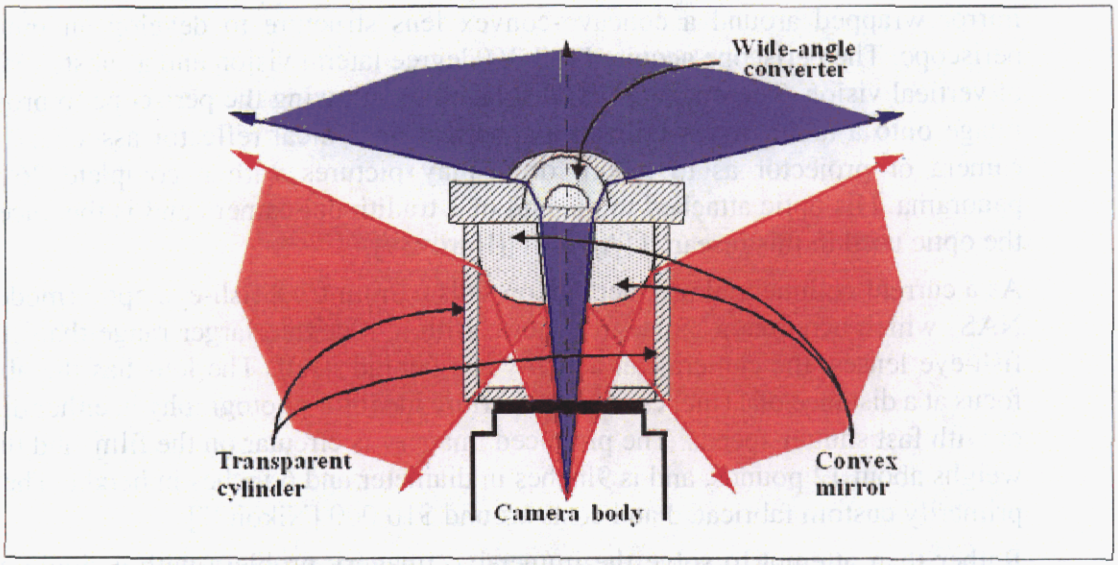


Figure 2-13: Schematic of panospheric optic for acquiring full immersive imagery

The term panospheric, trademarked by PVSI (Panoramic Viewing Systems, Inc.), refers to an image that contains more than a hemisphere ($>2\pi$ steradians) of visual field. In fact, it is possible to develop a single optic that can acquire a field of view in all directions simultaneously (full 4π steradian FOV). Panospheric imaging was pioneered in the Canadian Defense Research program as a possible solution for achieving visualization for crews of armored vehicles while in a "hatches-down" configuration. While not yet used for this application, DRES has demonstrated panospheric imaging as a passive observation post that enables remote surveillance.[Bogner95].

A combination of refractive and reflective optics is used to acquire a full panospheric image. Acquiring a hemisphere worth of imagery is achievable through fish-eye optics. Such an optic captures the hemisphere of imagery in the direction the camera is pointed. Likewise, taking a picture of a mirror provides an image in the direction behind the camera. The optic shown in Figure 2-13 shows how a combination of wide-angle optic and hemispheric mirror can be used to image virtually a full sphere worth of data simultaneously.

The basic effect of the hemispheric mirror is identical to the reflection achieved from the back of a spoon. The concept is an exaggeration of the curved mirrors commonly used for assisting with automobile blind spots or for intersections where visibility is difficult such as aisles in convenience stores not directly visible from the cashier position.

Related work in Panospheric Imaging

Panospheric and panoramic optics have been in existence for over 50 years, and numerous styles have been designed. Initial constructs were simple mirrors suspended in front of standard 35mm cameras. Alternative solutions involve extensive optic trains or multiple cameras. Recent developments take advantage are oriented toward exploiting advanced mirror designs. Numerous companies have developed formats for storing and dewarping panospheric imagery.

In the late 1930's, Wilbur Hardy and Waldemar Ayres designed perhaps the first panospheric optics, which are shown in Figure 2-14. Hardy utilized an inverted conic mirror wrapped around a concavo-convex lens structure to develop an innovative periscope. The periscope acquired full 360degree lateral vision and a substantial range of vertical vision. The imagery was displayed by allowing the periscope to project the image onto a table. Ayres utilized a spherical or conical reflector associated with a camera or projector as to record or display pictures with a complete 360degree panorama. His optic attached to the end of a traditional camera and is the ancestor of the optic used in this research [Ayres41][Hardy39].

As a current commercial product, Nikon sells a 6mm f/2.8 fish-eye optic (model 1405 NAS) which acquires a 220degree view. With a 40degree larger range than standard fish-eye lenses, the camera can actually see behind itself. The lens has the ability to focus at a distance of 11inches, and an aperture ideal for photography in either dim light or with fast shutter speeds. The produced imagery is circular on the film, and the optic weighs about 12 pounds, and is 9inches in diameter and 6 inches in height. The lens is primarily custom fabricated and costs around $10,000 [Nikon97].

Rather than attempt to solve the immersive imagery problem with a single camera, several options have utilized numerous cameras positioned close together, rather than spread out in an array. David McCutchen developed a dodecahedral system which could be used both for recording and projecting a full sphere of imagery. Vishvjit Nalwa at Bell Laboratories utilized an inverted pyramid and four cameras to create an image which could easily be remapped onto a cylindrical projection. Each reflected image had a 100degree field of view and the property of being acquired from the same focal point [McCutchen91][Nalwa96].

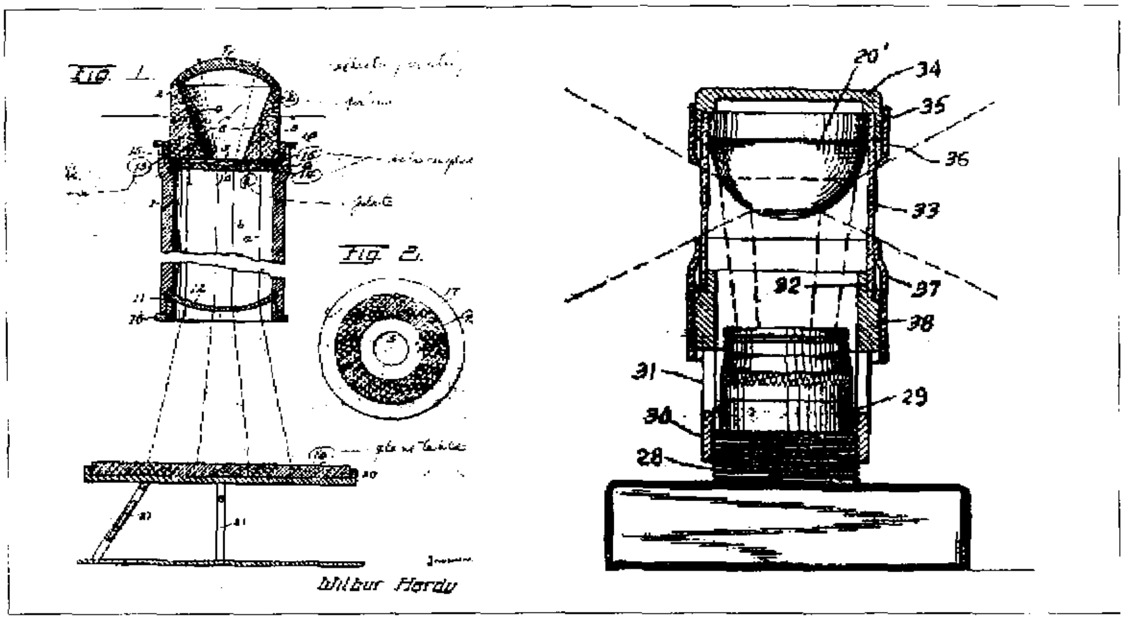


Figure 2-14: Precursors to the panospheric: the Wilbur Hardy periscope and the Waldemar Ayres film camera optic attachment as illustrated for their patent applications.

By playing with the format of the mirror, it is possible to alter the imagery acquired by a single camera. The Omnicam developed at the University of Columbia uses a mirror based on a parabolic shape. This results in an optic which has a single focal point for every part of the reflected image, and simplifies the processing required to generate an *undistorted perspective projection image*. To maintain this property, the field of view is limited to a hemisphere, and the mirror must be very well aligned axially with the taking optic. The Omnicam is intended mostly for teleconferencing or surveillance applications where the camera would normally be placed on a planar surface (table, ceiling, etc.) and is shown in Figure 2-15 [Nayar97].

PIER Corp, in association with the Canadian Defense Establishment - Suffield created a mirror format which creates a reflection that has a linear distribution of pixels over the entire reflective range. And PVSI and the University of Alberta have developed a dual lobed hemispheric mirror which places an axially aligned smaller radius hemisphere mirror on the surface of the larger mirror (like a snowman). This produces an image in which each location in the world is visible twice, and so can be used as a very-short vertical baseline stereo [Southwell96].

Companies such as IPIX (formerly Omniview) and BeHere, have developed their own proprietary standards and methods for acquiring and displaying still-frame panospheric imagery, and are developing video formats currently. Panospheric imagery has also been demonstrated using industry standards such as Quicktime VR and VRML [Kuban94].

Understanding Panospheric Images

An image from a full panospheric optic consists of two regions, the forward view and the backward view. The forward view is easy to comprehend, as fish-eye optics and wide-angle imagery is seen somewhat commonly. The image is circular in shape, with the point directly forward in the center of the image. Each radius corresponds to a

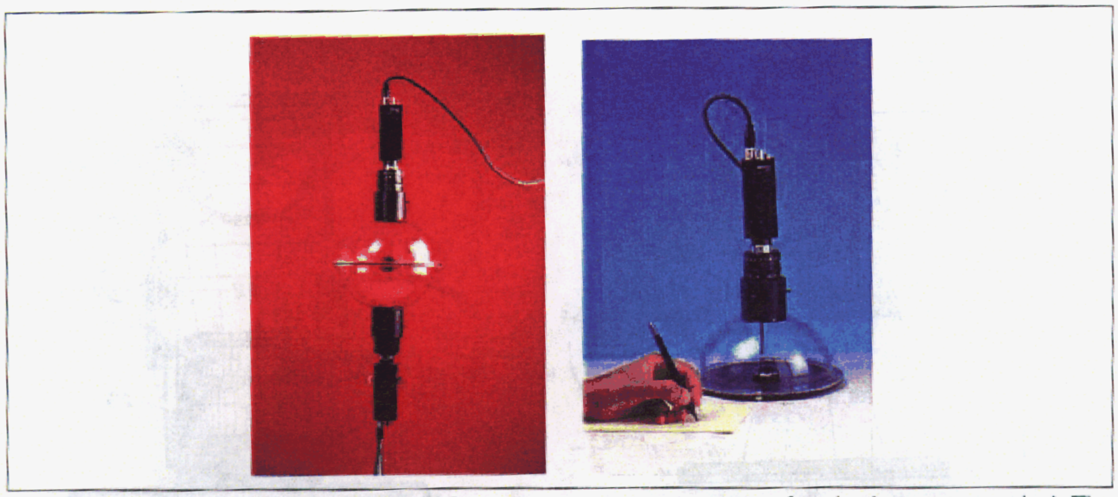


Figure 2-15: The Omnicam prototypes which possess a common center of projection at every pixel. The version on the left uses two cameras to achieve a full spherical view.

particular elevation angle, and the entire circle at that radius is the 360degree panorama about that elevation angle.

On the surface, the backward view (off the reflective optic) is nearly identical. The point in the center of the image is directly backwards (right through the center of the camera acquiring the image). Extending out in radius again corresponds to a particular elevation angle with the circle at that radius being the panoramic strip. In both views increasing in radius means moving in elevation angle closer to perpendicular to the direction the camera is pointing.

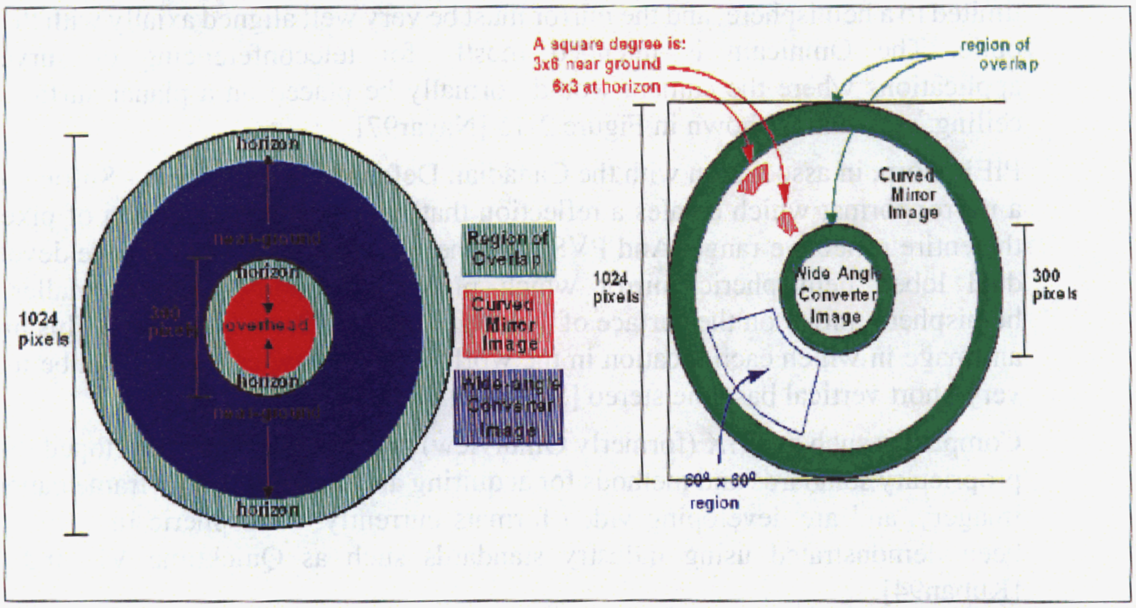


Figure 2-16: Schematics of panospheric images showing overlap between overhead and panoramic regions as well as the shape of square degrees at different radii.

When the two optics are combined, the central region of the backward view (that shows the reflection of the camera) is replaced with the entire forward view. As a result, starting from the center of the image, the angle described by increasing in radius starts at directly straight in front of the camera. While passing through the remainder of the forward view, the angle moves closer to looking sideways. At the transition from the forward to the backward view, the elevation angle jumps to pointing nearly straight backwards. Working through the remainder of the backward view, the elevation angle returns toward the perpendicular. At the edge of the backward view (the edge of the entire image), the view is liable to be overlapping with the outermost portion of the forward view. This is shown graphically in Figure 2-16.

The discontinuity which occurs at the transition between the two views and the resulting overlap in field of view that occurs at the outermost portions of each view provide for interesting processing. To create a seamless view these two regions need to be merged. This is easy in one respect since a given radial direction represents a particular azimuth angle. However, the number of pixels that comprise the entire panoramic strip differs based on the circumference of the circles at the overlapping radii. So the resolution of the forward view is always less than the resolution of the backward view.

Also interesting to note is that along with the horizontal resolution changing based on the circumference at the particular radius, the vertical resolution changes in a non-linear fashion. Due to the curved surface of the reflective mirror, the angle between two adjacent pixels increases as the radius increases for both the forward and backward views. This means that while the density of pixels per degree in the azimuthal direction is increasing, the density of pixels per degree in the elevation direction is decreasing. The result is that the total number of pixels per square degree remains roughly constant, but changes in shape from tall and skinny to short and fat.

Chapter Summary

- Imagery, especially peripheral views, plays a substantial part in achieving and maintaining situational awareness, and situational awareness has been demonstrated to be vital in teleoperation success.
- Human vision is the de-facto standard for displaying telepresence, although other animals do a better job of acquiring situational awareness. The technology to replicate human vision in an immersive environment is not yet available.
- Immersive video acquisition through the innovative panospheric optic is a viable solution to meet the goals of teleoperated exploration.

Chapter 3

Achieving Full-Surround Video

An imaging solution for **Achieving Full-Surround Video** for exploration is designed and simulated, with details given of the implementation and the display.

The design of the panospheric optic depends upon the desired fields of view, and the sensor mounted on Nomad to perform the research was based on a hemispheric mirror. This shape has advantages over other simple reflective optics, such as a conic or paraboloid. Models of the optic were created to determine the size and distance the mirror needed to be placed from the taking lens, and the model was simulated in a rendering environment to verify the desired field of view was achieved.

The sensor was implemented in iterations, with a proof of principal version preceding the design and simulation, and a prototype being created to enable the development of the display and processing algorithms. The optic, camera, housing, and electronics of the final version are described in detail. General specifications of the implemented sensor are presented.

As the raw video from the panospheric camera was awkward for direct use during telexploration, each image was dewarped in software and formatted for both monitor (flat-screen) and immersive theatre displays. Two different methods of dewarping were implemented: using a look-up table approach, and applying texture-mapping/virtual-reality techniques. The Telepresent Interface, built on the latter approach, was used in the Electric-Horizon immersive theatre for evaluating telexploration by panospheric video.

Design of a Panospheric Optic

Acquiring a full 360degree panorama with extended vertical field of view (-70 to +30degree in elevation) as motivated by the Atacama Desert Trek can be accomplished using one of the simplest panospheric optics: a hemispheric mirror. The position and size of the mirror with respect to the taking lens of the camera determines the field of view actually captured by the optic. A spreadsheet of constants and equations is given for quantifying the expected field of view. Several examples are included to demonstrate the dependence of field of view on mirror size and position.

Why to Use a Spherical Mirror

Any of the panospheric designs enable the acquisition of a full 360degree panorama. The principal differences are in the vertical field of view and in the distribution of pixels over that range. Other factors are the focal point of the optic and the simplicity of construction.

A conic mirror is the ancestor of the panospheric optic. With a conic mirror, the vertical field of view is virtually determined by the angle of the cone. To have a view of 30degrees above the horizon at the edge of the image, the angle of the cone could be no more than 120degrees (60degrees on each side of the vertical) to prevent occluding the desired field of view with the mirror itself. Additionally, with an angle of 120degrees, the light rays coming from the desired +30degree elevation would be parallel to the mirror (at the grazing angle) and thus not very desirable.

The minimum elevation which could be seen in the reflection of the conic mirror occurs at the center of the panospheric image, with the light rays bouncing off the tip of the cone. This translates into light rays being completely vertical after bouncing off the slope (60degrees off the vertical) of the mirror. The reflection of a light ray of a plane is symmetric with the angle of incidence, and so the incoming light must be coming from 30degrees below the horizon, which is 40degrees too short of the desired downward field of view. There is no way to gather that particular 100degrees of field of view using a conic mirror.

The problem is that to see the near ground, a horizontal mirror is good, whereas to see the horizon, a 45degree tilt to the mirror is more ideal, and to see above the horizon, the required mirror begins to tilt toward the vertical. Replacing the cone with a surface that is faceted with a sequence of flat mirrors results in something that remarkably resembles the 1970's disco-ball. Taking the concept to the mathematical limit, a spherical mirror allows for smooth transitions between every angle of the field of view.

Why Not to Use a Spherical Mirror

Spherical convex mirrors (referred to as hemispheric mirrors in the remainder of this document) are not exactly the right shape, as the focal point and the distribution of the pixels over the vertical field of view are not ideal. A ray of light reflected off a sphere has an apparent focal point halfway between the center of the sphere and the surface. Thus, for the desired field of view, there is an arc of focal points which are maximally separated by a distance of half the radius of the sphere. This means that the camera needs to have a depth of field greater than this range in order to keep the entire image in focus, and generally suggests a small aperture, which limits the usefulness of the camera in low lighting conditions. Utilizing a parabolic mirror results in a single focal point, rather than the "focal-arc", and eliminates the problem.

The distribution of pixels along the vertical field of view is also not ideal with the hemispheric mirror. The central region of the hemispheric mirror acts exactly like a horizontal mirror, and results in imaging of the taking lens of the camera. Taking a central core out of the hemispheric and collapsing the mirror would eliminate this problem, but makes the mirror more difficult to fabricate.

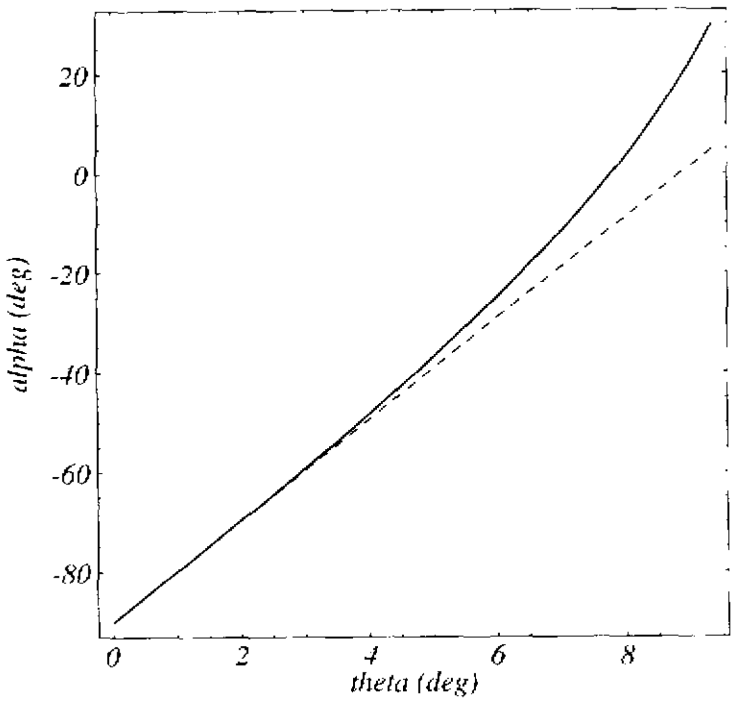


Figure 3-1: Mapping from angle of field of view (theta) to the reflected angle (alpha).

Regardless of the wasted central region of the mirror, the mapping of pixels from the radial line of the image to the real world vertical field of view is nonlinear. This means that more of the pixels are capturing reflections of the near ground than are capturing reflections of the horizon as shown in Figure 3-1. The effect is that achieving a particular vertical resolution at the horizon will result in significantly higher resolution being captured at the near ground. An optic can be designed so that this effect is eliminated. However, considering that the horizontal resolution is based on a circumference about the given radius, near-ground imagery has much less horizontal resolution than horizon imagery does, and so the hemispheric mirror in effect performs an inherent compensation, keeping the total pixels in every 2D region fairly equal.

The parabolic mirror as a solution for the focal problem of the hemispheric mirror has the disadvantage of having much worse pixel distribution problems at the angles above the horizon (due to the faster upswing of the parabola). And correspondingly, the mirrors that have equal pixel distribution do so by flattening the curve and result in a greater requirement for depth of field.

The hemispheric mirror thus has disadvantages, but a solution that fixes one disadvantage makes the other worse. In addition, hemispheric mirrors are common and fairly simple to create, unlike more advanced panospheric optics.

Constraining the Design

Having chosen to use a hemispheric mirror as the basis for the optic, the goal was to design the remainder of the optic to enable the desired field of view. Ideally, the optic would be as compact as possible, and utilize as much of the image space as possible. Creating the optic requires selecting the radius of the hemispheric mirror and the distance of the mirror from the taking lens. This needs to correspond to the field of view of the taking lens, and consider the potential occlusion of the downward field of view by the taking lens and camera.

As the size of the mirror decreases, the precision of the mirror grinding and the microscopic qualities of the surface begin to have a greater effect on the reflected image. Also, the smaller the mirror, the more the precision of the taking optic and its calibration have an effect.

As the mirror is moved further away from the taking lens, the relative size of the taking lens (and the camera) decreases from the mirror's point of view. This is important, as there is a goal for the sensor to acquire a useful view of 70degrees below the horizon. This constraint gives the maximum amount of the field of view that can be taken up by the taking lens and camera.

The maximum use of the image space is obtained when the image of the hemispheric mirror is inscribed on the CCD. More specifically, the chord of the hemisphere that provides the +40degree view should be inscribed. Theoretically, the field of view of the taking lens can be selected to make this possible for any sized hemisphere at any distance from the taking lens. However, most taking lenses have a minimum distance of focus, and so there is a constraint on how close the mirror can be to the taking lens.

Before beginning to design the optic, a taking lens was selected. The intent in selecting the taking lens was to have a known minimum focus-distance to set a lower bound on the distance from the mirror to the taking lens. This enabled sizing of the mirror to ensure that the taking lens did not interfere with the desired downward viewing angle. While microscope optics can focus very closely, they have a very short focal range and a very tight field of view which would create the undesirable need for a very small mirror located close to the taking lens. A standard 35mm F-mount macro-zoom taking lens was selected, as it had extended focal range, and ability to adjust the field of view. The minimum focal distance was around 5 inches from the end of the optic (depending on the amount of zoom). As the size of the CCD in the camera was not known, the ability to zoom allowed the design to compensate for the apparent change in field of view between a 1/2inch and 1/3inch CCD.

Sizing the Mirror

A spreadsheet was created in Mathematica to allow evaluation of various sizes of mirrors located at various distances from the taking lens. The spreadsheet started with a set of constants that described the physical characteristics of the taking lens and camera as well as the minimal focal distance and field of view of the taking lens. The distance from the taking lens that resulted in inscribing the chord representing the +30degree field of view in the cone specified by the field of view was then calculated. At that distance, the downward viewing angle at the point of occlusion was then calculated. From the spreadsheet, a mirror radius and distance from the taking lens were obtained.

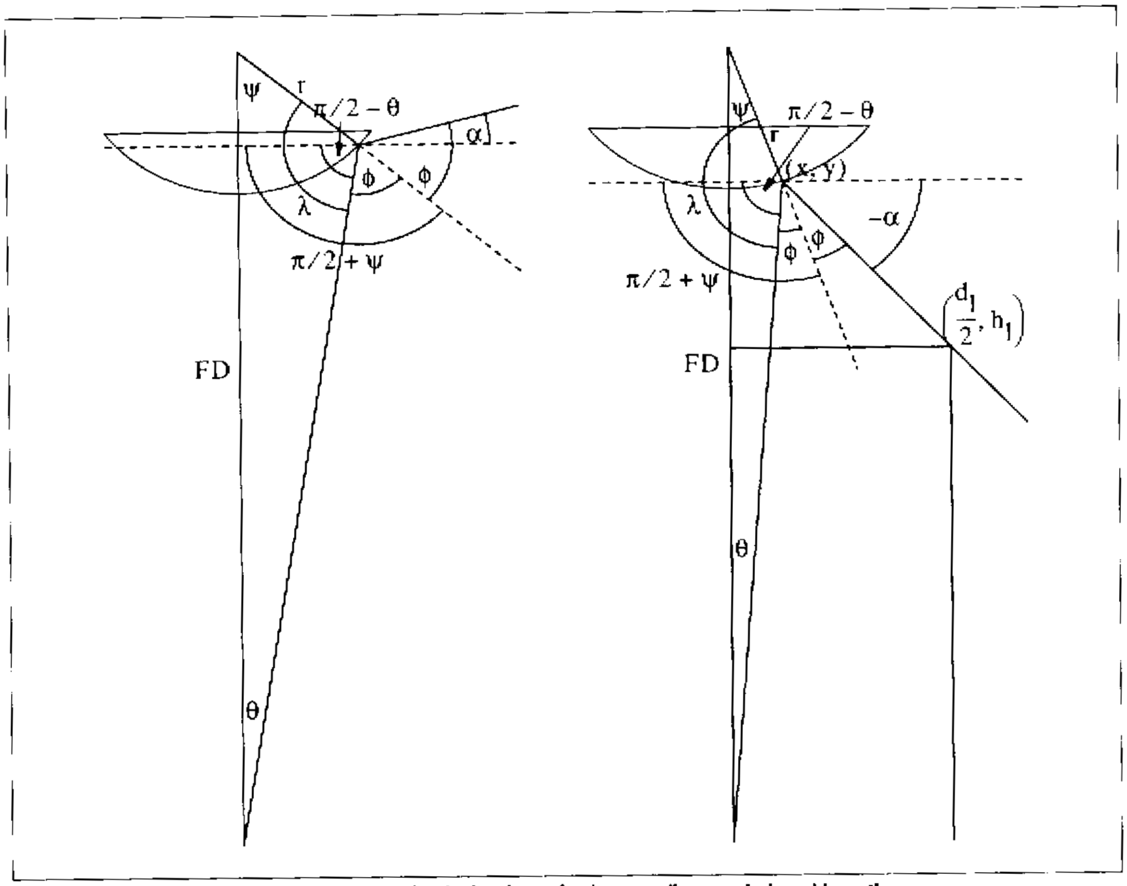


Figure 3-2: Description of variables used in derivation of mirror radius and chord length

Given the focal distance FD, field-of-view FOV, and a maximum up-angle α_{max}, the required mirror geometry is calculated, given by the radius r and the chord length c of the spherical mirror. Using the geometry shown in Figure 3-2,

$$\left(\frac{\pi}{2}-\theta\right)+2\phi-\alpha = \pi \quad \text{and} \quad \left(\frac{\pi}{2}+\psi\right)+\phi-\alpha = \pi \quad . \qquad \text{Eqn. [3-1]}$$

$$\text{Thus,} \quad \phi = \frac{1}{2}\left(\frac{\pi}{2}+\alpha+\theta\right) \quad \text{and} \quad \psi = \left(\frac{\pi}{2}+\alpha-\phi\right) \quad . \qquad \text{Eqn. [3-2]}$$

$$\text{Combining these two equations} \quad \psi = \left(\frac{\pi}{4}+\frac{\alpha}{2}-\frac{\theta}{2}\right) \quad . \qquad \text{Eqn. [3-3]}$$

$$\text{Now,} \quad \frac{\sin(\lambda)}{FD} = \frac{\sin(\theta)}{r} \qquad r = FD\frac{\sin(\theta)}{\sin(\lambda)} \qquad \lambda = \pi-\theta-\psi \quad \text{and}$$

$$r = FD\frac{\sin(\theta)}{\sin(\pi-\theta-\psi)} = FD\frac{\sin(\theta)}{\sin(3\pi/4-\alpha/2-\theta/2)} \qquad \text{Eqn. [3-4]}$$

$$\text{Hence, for} \quad \theta_{max} = \frac{1}{2}FOV \quad , \quad \alpha_{max} \qquad \text{Eqn. [3-5]}$$

$$r = FD\frac{\sin(\theta_{max})}{\sin(3\pi/4 - \alpha_{max}/2 - \theta_{max}/2)} \quad , \quad c = 2r\sin(\psi) \quad . \qquad \text{Eqn. [3-6]}$$

Given the mirror radius r, the focal distance FD, and the ray angle θ the reflection angle α is calculated:

$$\frac{\sin(\lambda)}{FD} = \frac{\sin(\theta)}{r} \qquad \text{Eqn. [3-7]}$$

$$\sin(\lambda) = \frac{FD}{r}\sin(\theta) \qquad \text{Eqn. [3-8]}$$

For our setup, $\frac{\pi}{2} < \lambda < \pi$, such that $\lambda = \pi - \sin^{-1}\left[\frac{FD}{r}\sin(\theta)\right] = \pi - \theta - \psi$ and

$$\psi = \sin^{-1}\left[\frac{FD}{r}\sin(\theta)\right] - \theta \qquad \text{Eqn. [3-9]}$$

From Eqn. [3-1], $\alpha = 2\psi + \theta - \frac{\pi}{2}$ and so $\alpha = 2\sin^{-1}\left[\frac{FD}{r}\sin(\theta)\right] - \theta - \frac{\pi}{2}$ Eqn. [3-10]

Given an angle α, Eqn. [3-10] Eqn. [3-10]can also be solved numerically for θ.

Given a field-of-view $FOV = 2\theta_{max}$, maximum up angle α_{max}, and maximum down angle α_{min}, and the lens diameter d_l and lens height h_l, the minimum required focal distance FD and mirror geometry (r, c) are calculated.

First, note that $r = kFD$, where $k = \frac{\sin(\theta_{max})}{\sin(3\pi/4 - \alpha_{max}/2 - \theta_{max}/2)}$ = constant Eqn. [3-11]

Eqn. [3-10]From Eqn. [3-10], $\alpha = 2\sin^{-1}\left[\frac{1}{k}\sin(\theta)\right] - \theta - \frac{\pi}{2}$ Eqn. [3-12]

and $\alpha_{min} = 2\sin^{-1}\left[\frac{1}{k}\sin(\theta)\right] - \theta - \frac{\pi}{2}$ Eqn. [3-13]

Eqn. [3-13] Eqn. [3-13] can be solved numerically for θ. The equation of the line intersecting the lens body is given by:

$$(y - h_l)\sin(\psi + \phi) = \left(x - \frac{d_l}{2}\right)\cos(\psi + \phi) \qquad \text{Eqn. [3-14]}$$

From Eqn. [3-1], $\phi = \psi + \theta$, thus $(y - h_l)\sin(2\psi + \theta) = \left(x - \frac{d_l}{2}\right)\cos(2\psi + \theta)$ Eqn. [3-15]

Also, the coordinates (x,y) of the reflection point on the mirror surface are given as:

$x = r\sin(\psi) = kFD\sin(\psi)$ and $y = FD - r\cos(\psi) = FD[1 - k\cos(\psi)]$. Eqn. [3-16]

Combining Eqn. [3-15] and Eqn. [3-16]:

$$[FD[1 - k\cos(\psi)] - h_1]\sin(2\psi + \theta) = \left[kFD\sin(\psi) - \frac{d_1}{2}\right]\cos(2\psi + \theta) \qquad \text{Eqn. [3-17]}$$

$$FD = \frac{h_1\sin(2\psi + \theta) - \frac{d_1}{2}\cos(2\psi + \theta)}{[1 - k\cos(\psi)]\sin(2\psi + \theta) - k\sin(\psi)\cos(2\psi + \theta)} \qquad \text{Eqn. [3-18]}$$

where $\psi = \sin^{-1}\left[\frac{FD}{r}\sin(\theta)\right] - \theta = \sin^{-1}\left[\frac{1}{k}\sin(\theta)\right] - \theta$ and Eqn. [3-19]

$$k = \frac{\sin(\theta_{max})}{\sin(3\pi/4 - \alpha_{max}/2 - \theta_{max}/2)}$$ and θ is the numerical solution to Eqn. [3-13] Eqn. [3-20]

The mirror geometry (r, c) can now be determined directly from Eqn. [3-6]Eqn. [3-6].

Effects of Changes in the Design

To demonstrate the effect on the optic of adjusting the size of the mirror, the field of view of the taking lens, and the size of the camera body several cases are shown below. The first case (baseline) is the one identified as a successful possibility, and from which the sensor prototypes and software processing were derived.

example 1: baseline

The mirror sized by the above equations to produce a +30, -70degree field of view with a minimal focal distance constraint of 8inches and a field of view of 18.7degrees was 32.6mm in radius and a chord length of 60mm was utilized.

example 2: mirror resize

With a smaller mirror, the only way to inscribe the appropriate chord length in the field of view is to bring the mirror closer to the taking lens. This results in severe obstruction of the downward field of view. It also violates the constraint imposed by the minimal focal distance of the taking lens

example 3: field of view adjustment

Increasing the field of view means that the inscribed chord length is longer, and so the mirror must be appropriately increased in radius. The effect is to bring the front surface of the mirror closer to the taking lens. As the mirror gets closer, the lens begins to obstruct more of the downward field of view, and takes up a relatively larger region of the reflected image.

example 4: camera/lens size change

If the camera parameters change in physical size (such as the taking lens resize shown here), the result is that the downward field of view is no longer attainable due to the obstruction of the lens. To solve this problem, the mirror must be moved further from the taking lens, which requires either an increasing of mirror radius, or a decreasing of field of view. While there is only one solution for the mirror size at the minimal focal distance with a particular field of view, there is actually a family of solutions which can create an optic that acquires the desired field of view.

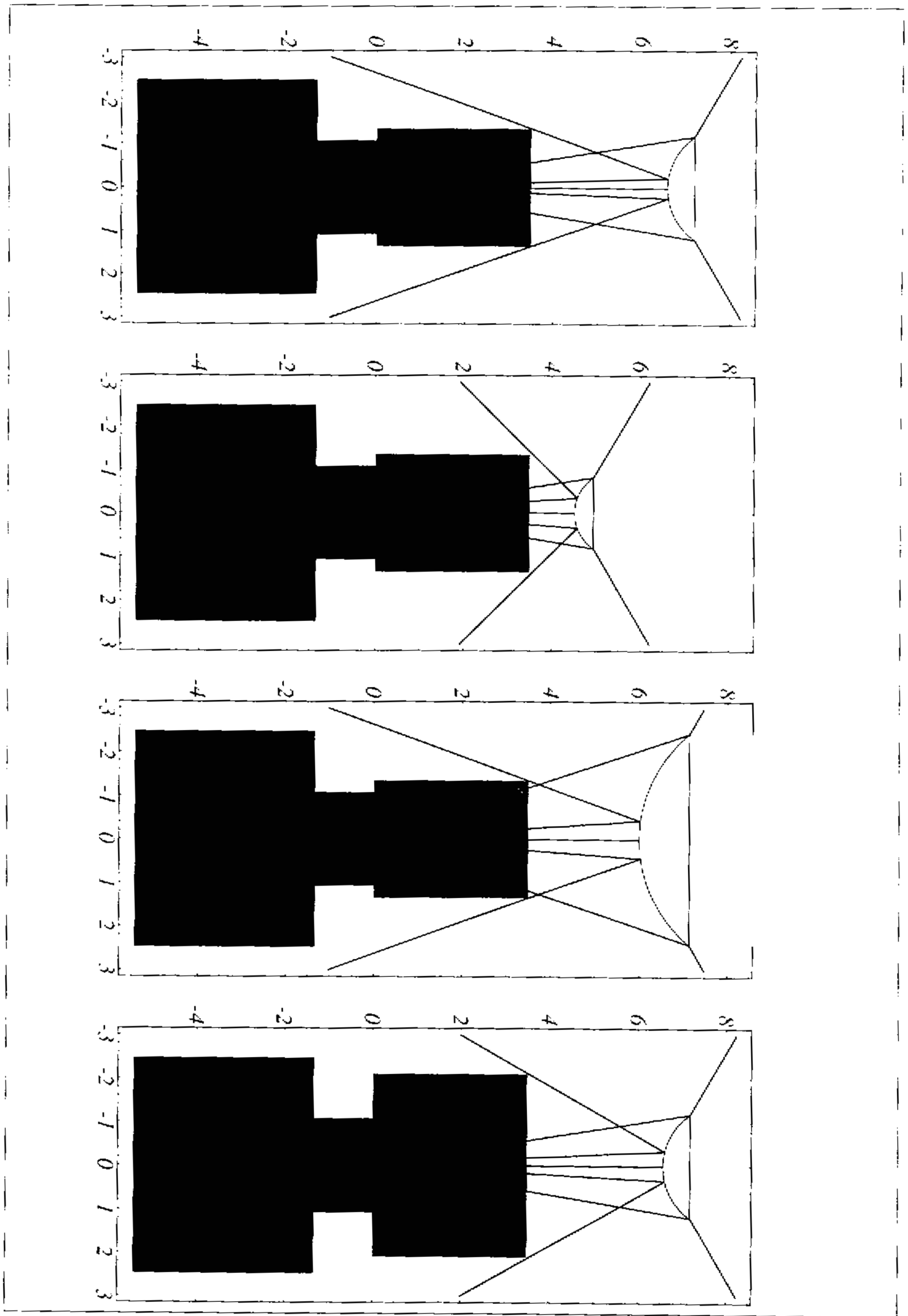


Figure 3-3: Drawings of the four examples showing effects of changes in the optic design. baseline, smaller mirror and constant field of view, wider field of view and constant focal distance, larger taking lens

Simulation of the Optic

To evaluate the designed optic, confirm the field of view, and create sample imagery needed for developing the processing and display algorithms, a rendering environment was implemented. A geometric model of the optic was created using the values resulting from the design. Recursive raytracing techniques were used to trace the path of light rays incident on the reflective surface. Two scenes were created into which the optic was placed, and panospheric imagery was acquired. From the same points in the world environment, several standard field of view images were rendered to allow for comparison when the panospheric imagery was dewarped and processed later.

Still frames were rendered using recursive raytracing techniques. The environment allowed for point light sources such as the sun, and creation of shadow rays. The available shape primitives were spheres and polygons, and the environment was capable of implementing recursive reflection and refraction rays, and so could correctly model the hemispheric mirror. Also included as features were light source attenuation, to allow for modeling of night operations with local illumination sources, automatic exposure adjustment, fog effects, and either bounding box hierarchies or voxel quadtree subdivision of the scene (principally for improved rendering speed).

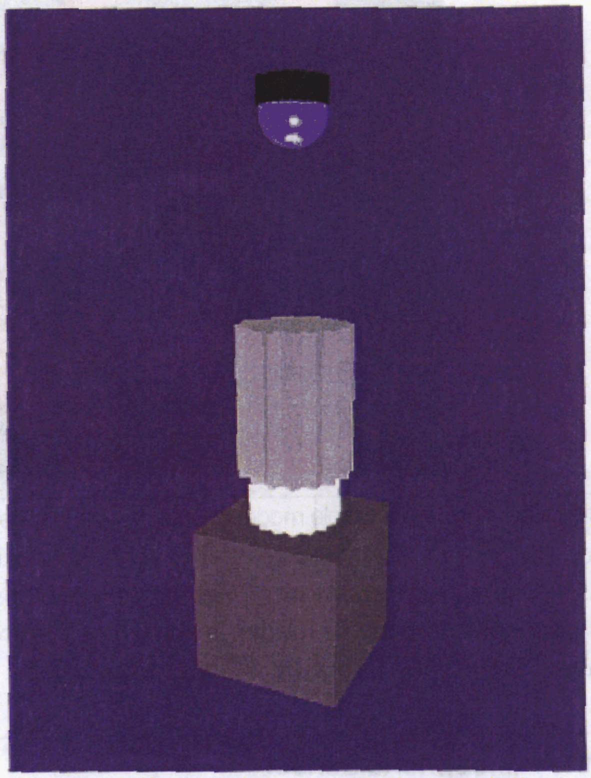


Figure 3-4: Rendered model of the panospheric camera

The first world model created was of an indoor scene, with a table placed in a square room so that understanding the correctness of the dewarped image would be easier. The second world model was of an outdoor environment, with ravines and dunes, with the camera sitting on top of a mock robot body. This model was rendered both in "daylight" and in night, with faint ambient illumination and a light source on the robot.

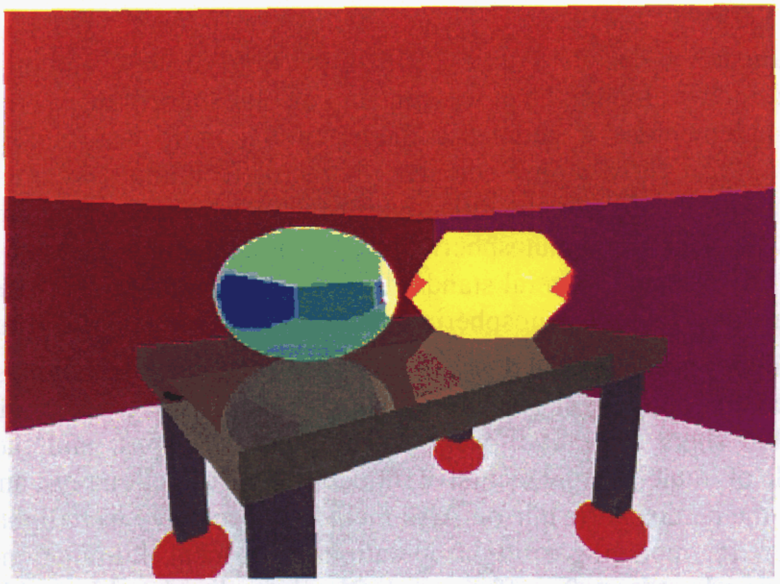


Figure 3-5: Rendered view of indoor table scene, using a 80degree FOV and perspective projection

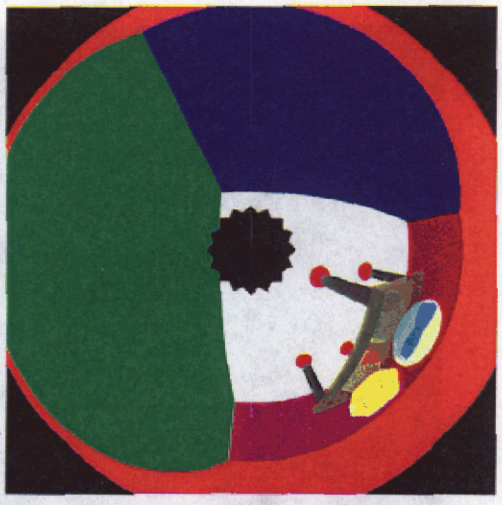


Figure 3-6: Rendered view of the indoor table model through the panospheric optic (black star is taking lens)

The results of the simulation were to verify that the designed optic did acquire the desired fields of view, and creation of the prototype sensor was begun. A rendered image of the panospheric camera model is shown in Figure 3-4 and sample images of the indoor model are shown in Figure 3-5 and Figure 3-6.

Implementation of the Panospheric Sensor

An initial proof of principal version of the sensor was created to demonstrate the capabilities of panospheric imaging in general to the Atacama Desert Trek design team. After the above design and simulation processes were performed, a prototype sensor was created and used to develop algorithms for dewarping and display. After working with the prototype, a refined and miniaturized optic was created and packaging was designed and developed to help weatherproof and ruggedize the sensor. No attempt was made to miniaturize the electronics of the sensor.

Proof of Principal Version

To convince designers on the Atacama Desert Trek team that panospheric imaging could elegantly generate the desired field of view, a proof of principal version of the panospheric sensor was constructed. A 35mm camera acquired still images taken of the reflection on a polished aluminum hemispheric mirror. The mirror was cored so that it could be fitted with wide-angle lensing to serve as a proof of principal for the full immersive image acquisition as well as just the extended panorama.

The hemispheric mirror was suspended above the camera using three rods positioned evenly about the mirror circumference. The rods were threaded to allow adjustment to the height of the mirror. The acquired image did include these suspension rods, and each one occluded roughly 5degrees of the panorama. The mirror was approximately 160mm in diameter, and had a 15mm lip into which the rods were attached. This lip also extended into the field of view of the acquired image, occluding the top region of the vertical field of view.

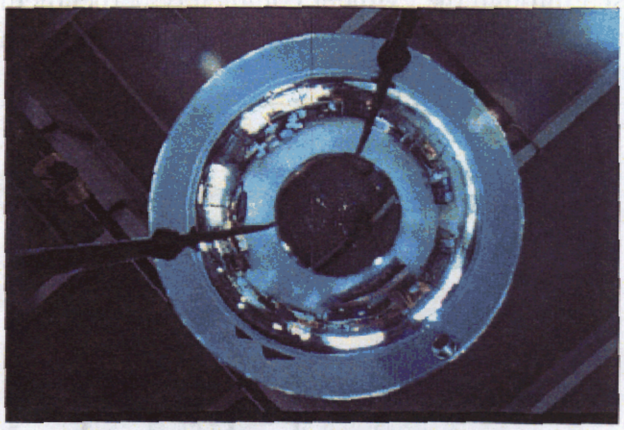


Figure 3-7: Raw image obtained from the 35mm camera using the proof of principal optic

The center 50mm of the hemispheric mirror was cored and the back side of the mirror was hollowed out to allow placement of a wide-angle lens which could capture the top side hemisphere needed to produce a full panospheric optic. However, the focal point of the wide angle lens was at infinity by the nature of the lensing. This required that the camera have a depth of field to image both the mirror (which was about 25cm from the lens) and infinity and in general was unsuccessful. As additional effort would be required to construct a suitable wide angle lens for embedding in the back side of the hemispheric mirror, and was not appropriate for the Atacama Desert Trek, this portion of the proof of principal version of the panospheric optic was declared to be a future extension.

The images were captured on slide film, and digitally scanned into the computer. Because the optic had not been ideally designed, the reflected image was less than inscribed, and a bit far away from the taking lens, making the cored region from the center a little too large. In addition, no processing or dewarping of the imagery was performed. However, the imagery was sufficient to demonstrate that panospheric imaging would be appropriate for the Atacama Desert Trek, and the modeling and simulation work described above was begun, with the result leading to the prototype sensor.

Prototype Sensor

To create the prototype mirror, a plano convex lens was mirrored with a first surface AlSiO coating. Rather than using a full hemisphere, a 100mm diameter chord of a 120mm focal length lens was selected, as the upper part of the hemisphere added information beyond the +30degree vertical horizon, and so was not needed. The focal length takes into account the refraction of the light through the glass (1.523 index of refraction), and so the actual physical radius of curvature is 62.76mm.

The three suspension rods from the proof of principal version were replaced with an acrylic cylinder that was 152.5mm in diameter, with a 6.35mm thickness. The mirror was suspended inside the cylinder on an adjustable rod. The cylinder extended down around the entire taking lens, and attached to a platform in the plane of the focal point, as shown in Figure 3-8. The acrylic cylinder did add distortion to the imagery, as the cylinder varied slightly in thickness and density. Also, imagery of the near ground passed through the cylinder at a 70 degree angle of incidence, while imagery of the horizon passed through the cylinder completely perpendicular. These distortions were noticeable in the imagery, but not severe enough to diminish the effect of the sensor as a working prototype.

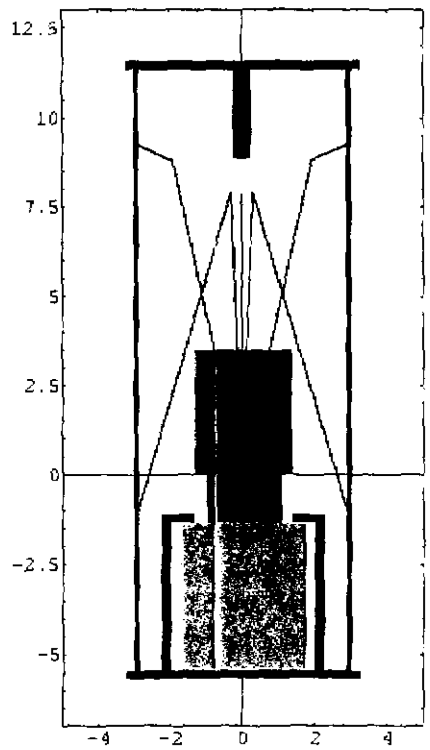


Figure 3-8: Schematic of the prototype camera and optic assembly

The camera, produced by DALSA Inc. as model CA-D2, acquired 8 bit grayscale digital imagery at a resolution of 512 x 512 pixels. The camera was connected to a

90Mhz Pentium PC running Windows NT 4.0 via a Bitflow Data Raptor framegrabber. Acquisition of raw imagery could be performed at 10Hz. Utilizing a perspective projection dewarping algorithm based on look up tables, the imagery was displayed on the computer monitor at an update rate of 0.2Hz.

Figure 3-9: Image acquired using prototype camera and optic assembly

Full Implementation of Sensor

The intent of the full sensor implementation was to miniaturize and refine the optic, upgrade the camera, ruggedize the assembly, and perform compression and transmission of the video stream.

Refining the Optic

While the taking lens remained constant from the prototype to the final version, the mirror and enclosure were modified. In addition, a shield was added to the top of the sensor to black out useless parts of the imagery.

The hemispheric mirror remained a first-surface mirror-coated plano-convex lens, but was reduced in size from the 120mm focal length and 100mm chord diameter to a 70mm focal length and 60mm chord diameter. This resulted in a 36.61mm radius of curvature on the mirror. It was again suspended via an adjustable rod, which after being focussed appropriately was locked down for the duration of its use on board Nomad.

The taking lens remained the F-mount Nikon macro-zoom lens. Model number 1959NCP, the lens is a 28-85mm f/3.5-4.5 autofocus zoom, with a minimum focal distance of 250mm. Due to the reduced diameter of the mirror, and an increased size of the CCD, the field of view of the taking lens had to be significantly reduced, and so the length of the taking lens was increased over the prototype.

Replacing the acrylic cylinder was a 150mm long Corning Pyrex 7740 medium-wall glass tubing with a 75mm outer diameter, and a 2.4mm thickness. This cylinder was epoxied on both ends with 62mm to 77mm stepper rings. One end of the cylinder was attached directly to the Nikon taking lens, and the other was fitted with a metal lens cap. A threaded rod was then inserted through the center of the metal lens cap to allow the mirror to be suspended inside the cylinder at the appropriate height.

Above the metal lens cap was placed a circular disk of sandblasted and black anodized 1/8" thick aluminum. As the reflection of the mirrored surface is an inscribed circle, the corners of the square image in the circular region are acquiring useless imagery. The disk served to effectively black out these corner regions so that the image consisted only of useful data. This had the advantages of eliminating any direct overhead information that might saturate the CCD, and enabling compression of the entire raw image rather than requiring a processing step to black out the unwanted regions.

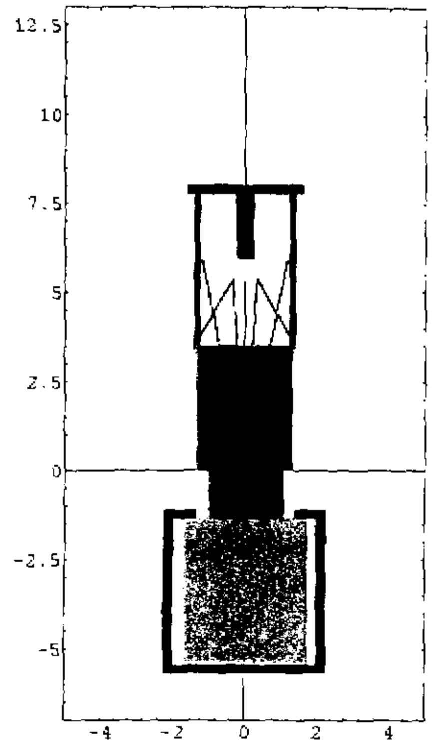


Figure 3-10: Schematic of the final version of the panospheric optic and camera assembly

The refinement of the sensor reduced the size of the entire assembly from 300mm to 75mm in diameter, and 900mm to 300mm in length. The reflective surface was reduced to about 60% of its original size, both in diameter and radius. The assembly also allowed direct access to the taking lens (which before was encased in the cylinder), making fine adjustments of focus and exposure possible. Once the assembly was focused and adjusted for the expected illumination of the Atacama Desert, it was locked down and sealed for the entire duration of the trek.

Upgrading the Camera

The intent in upgrading the camera was to add some form of color acquisition, and to increase the pixel depth at each point to help with setting the exposure. The DALSA CA-D2 was replaced with the CA-D4/7 which provided 12 bit grayscale and a resolution of 1024 x 1024 pixels. The digital output was sent via RS-422, and the camera was capable of acquisition at 8 frames per second. The CCD had a pitch of 10micron, 100% fill, and operated in a full frame transfer mode. The camera was capable of acquiring imagery at about 8 frames per second at the fastest exposure settings, and about 6 frames per second at slow settings used in low lighting conditions.

To achieve color, a RGB filter was directly applied to the CCD, so that each quartet of pixels contained two green pixels, one red pixel, and one blue pixel. As this was a new procedure for the camera manufacturer, there was no way to adjust the gains of the various colors, and so the white balance of the camera was off, with the image of a white target appearing a bit red. To compensate for this, the bits of the red pixel were shifted once. Pixels that had been saturated were left as saturated, and occasionally the images would be unpalatable as a result. During the Atacama Desert Trek, the white balance feature was toggled, depending on which setting looked better for the current conditions. To complete the colorization, every four pixels (48bits of data) was transformed into a single RGB triple (24bits). The two green values were averaged together to give a single green value, and the 12 bits at each pixel were converted to 8bits. The resulting 512 x 512 24bit image was sent to the compression engine.

As it turned out, the desert was significantly brighter than anticipated and there were virtually no shadows during the bulk of the daytime operations. Thus the increased resolution was not as important as it had been in the cloudy field testing in Pittsburgh. For the most part, to prune from 12 bits to 8, the highest order 8 were kept, with the low order bits ignored. At dawn and dusk, when there was limited light, the highest order bits were ignored, and the middle eight of the twelve were used as a form of illumination enhancement. This created rather fuzzy images due to the compounding of compression and noise effects of those low order bits, but the shifted configuration did allow for extended operation hours.

Figure 3-11: Raw image from the final camera and optic assembly when mounted on the robot.

Ruggedizing the Optic Assembly

No attention was paid to ruggedizing or packaging the prototype optic, as it was purely intended for desktop use in the lab, with occasional placement outdoors or on the robot to check field of view and exposure. The final version needed to be secured against vibration, water, and dust.

The camera and taking lens were enclosed in a torpedo shaped structure of aluminum. The bottom portion of the torpedo was left open for cables, and packed with foam rubber to seal against dust. The top of the torpedo fit tightly around the glass cylinder with a rubber O-ring serving to make a waterproof seal. The torpedo had mounting points to allow the camera to be bolted to the sensor mast, and clamped down tightly around the camera itself, serving both to restrict vibration and to serve as a heat sink for the camera.

The metal lens cap at the top of the camera was replaced with a fitted aluminum cap that was again sealed to the glass cylinder with an O-ring. The shield disc was overlaid on the cap, and a bolt tightened down on the top to hold suspended mirror and the disc in place. A conic top piece was added to cover the bolt and seal down snugly on the shield disk to prevent dust or moisture from working into the mirror enclosure via the same hole as the threaded rod which held the mirror and the cylinder cap.

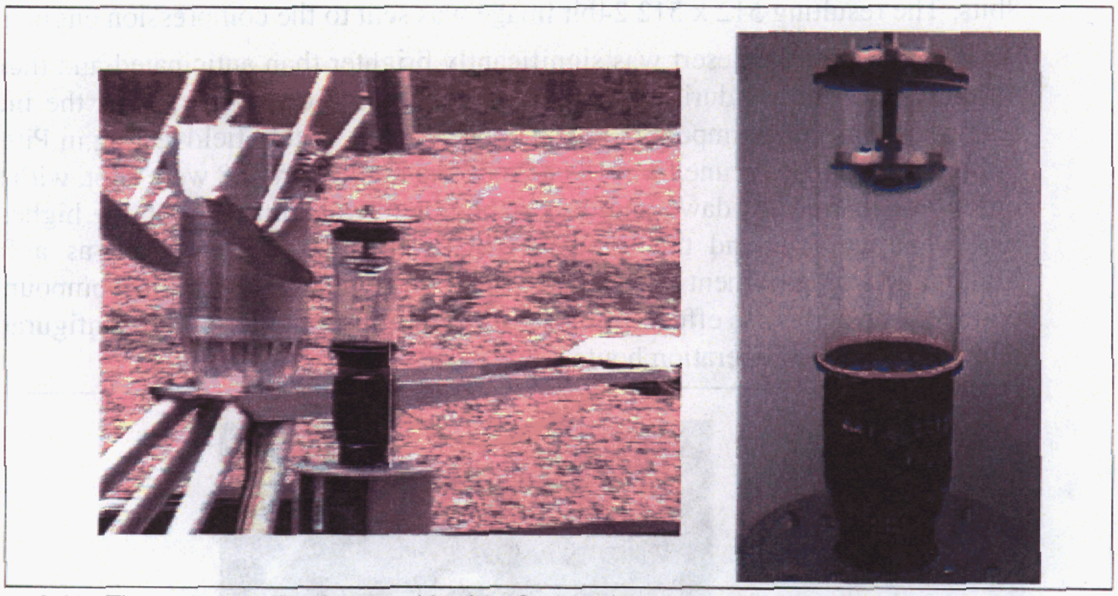


Figure 3-12: The camera and optic assembly of the final version, shown without environmental casing.

The remainder of the electronics of the sensor was shock mounted inside the electronics enclosure of the robot, and the cables were screwed down tightly and locked at each connection point to reduce failures. At the bulkhead between the electronics enclosure and the outside of the robot, the bundled cable for the sensor was cut in half and fitted with milspec connectors to enable a simple disconnect of the optic assembly in case a need arose to remove it for servicing while in the desert.

Acquiring, Compressing, and Transmitting the Video

As in the prototype, a standard desktop computer served to do the on-board processing, and a dedicated power supply was used for creating the required voltages for the camera. Acquisition, compression, and transmission of the imagery were accomplished using commercial products.

Unlike the prototype which needed to only perform look-up table transformations for displaying the imagery, the deployed sensor was required to perform real time compression, to buffer and packetize the video for transmission, and to perform pre-processing and control of the acquisition process. A 200Mhz dual Pentium-Pro processor desktop computer was selected. Running the multi-processor Windows NT 4.0 kernel, one processor was dedicated to performing compression, while the other handled the pre-processing, communication, and control.

To provide the power and appropriate voltages to the camera, a dedicated power supply was used (model PS-4 from Vision1 Inc.). The power supply ran off the AC power bus created by the on-board generator, and provided the +/-5V, -12V, and +24V needed for the camera logic, and charge gradient for the pixels.

The Bitflow series of framegrabbers was kept, but the Raptor was replaced with the Roadrunner due to the reduced form factor and the ability to use dual ported memory and reduce cycles on the processor for copying image frames. The Roadrunner also had drivers for the DALSA CA-D4/7 camera, which allowed for easy adjustment of the exposure duration.

Compression was accomplished using a wavelet-based compression software package from Summus Ltd. Compressing a video frame required about 22milliseconds, meaning that four frames could be realistically transmitted every second. The targeted video bandwidth for the communication subsystem off Nomad was 1.0megabit/second. With each raw frame of video using 6.3megabits(512x512@24), and a framerate of 4 per second, a ratio of 25:1 compression was required. After using the sensor in the desert, it was decided to sacrifice some of the framerate to increase resolution, and so the general transmission speed was 3Hz.

A FIFO buffering strategy was used to packetize the imagery, which was then broadcast using multicasted UDP packets to the control stations in both Pittsburgh, NASA Ames, and Santiago, Chile. The NDDS software package version 1.11d developed by RealTime Innovations Inc. was used to perform the transmission and reception.

Displaying Panospheric Imagery

Depending upon the method of display, dewarping is needed to display panospheric imagery in a format familiar to the human eye. Two methods for performing dewarping are to use a look-up table of per-pixel transformations and to import the raw image as a texture into a virtual-reality style domain. The first method was used to evaluate the design of the optic and for presentation in a format comparable to traditional imaging. A version of the latter method, the Telepresent Interface, was implemented as part of the display in the Electric Horizon theatre, the control center for the Atacama Desert Trek.

Methods of Display

Able to be formatted for display on domed theatres or for television, panospheric video is exceptionally versatile. Imagery for domes or wraparound screens must typically be a seamless fusion of multiple projectors. Panospheric imagery can be formatted for these type of displays, and theoretically the optic itself could be used for projection of the video from a single source. Panospheric imaging can just as easily be displayed for a flat screen, such as television, computer monitor, or printout. Alternatively, the image stream can be displayed and manipulated interactively using virtual reality techniques.

The most common technique for displaying panoramic imagery is to use multiple projectors that each are responsible for filling some fraction of the display. The original panoramic displays were actually polygons of flat screens, and the imagery either displayed slight parallax at the seams, or had to be acquired using lensing that distorted the image on the film so that it would appear correct, such as cylindrical lensing. This meant that the imagery was not very portable, as it could only be displayed in theatres designed with identical specifications. OmniMax movies continue to use this paradigm, with a multiple projector setup that is replicated identically in each theatre.

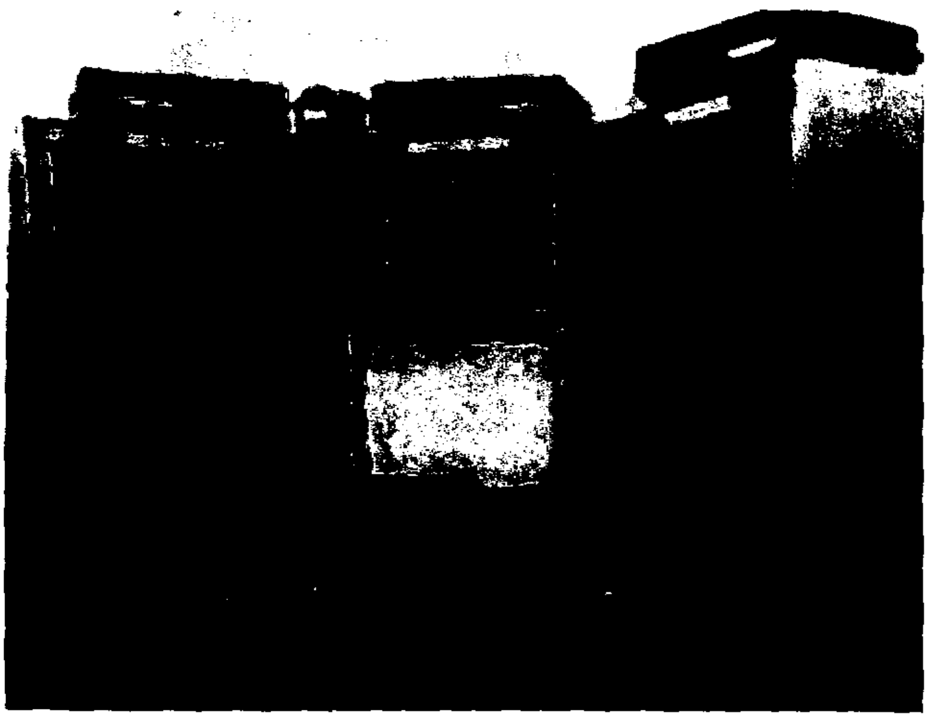


Figure 3-13: Multiple projector setup used at Sandia National Laboratory during McGovern's late 1980's teleoperation studies

Alternatively to using multiple projectors, if the intended display surface is spherical, or a subsection of a sphere, an identical panospheric optic to the one that acquired the imagery can be used as a projection source. With the optic positioned at the center of the sphere defined by the display surface, the video could be projected into the optic as though the optic was the target display. The resulting reflection, like the mirrored disco-ball, would distribute the imagery over the display. Those subsections of the sphere that are not desired for display would need to be blocked either by hardware or software means. While never implemented during this research, the concept remains intriguing.

Instead of constructing the display surface and the projection equipment to fit the acquired video, image streams from a panospheric sensor can be tailored to any display. To convert the panospheric video into another format entails creating a mapping of pixels from the raw image to the desired format. While cylindrical or spherical projections are found in immersive displays, the most common display surface is the flat-screen, used on televisions, movies, computer monitors, paper printouts and film prints. These flat screen formats take advantage of perspective projection to make the images appear normal to human vision. A sample method for mapping from panospheric to perspective projection is described below, but the process can be customized for any display.

Rather than considering the imagery as a static source for projection in one single format onto a particular display, virtual reality techniques can be used to allow a more interactive display experience. Starting with the panospheric image applied as a texture to a sphere, a virtual camera can be added to the scene and render some view. When the camera is placed at the center of the sphere, the image generated by the virtual camera is identical to the image which would be acquired using a real camera to take a picture of that direction from the sensor position in the real world. Multiple virtual cameras can be used, and they can be zoomed or rotated to create additional effects. This approach was used by the Telepresent Interface described below.

Perspective Projection Dewarping

As the geometry of the panospheric optic is constant, a mapping can be determined to dewarp the raw image and convert it into any other format. The process of image dewarping was developed commercially for a fish-eye lens by the Omniview product originally created by TeleRobotics International, Inc. Steve Bogner of the Canadian Defence Research Establishment - Suffield, who champions panospheric imaging for its use in armored vehicle visualization, claims that even a reflection off the "kitchen sink" can be dewarped and provide a normal view. Numerous techniques have been demonstrated for image dewarping, including several firmware implementation. The only requirement is having a model of the optic or reflective surface that acquired the imagery [Bogner95][Juday91][Kuban94].

A method of dewarping panospheric images was implemented in this research to verify the field of view of the designed optic, and to enable display of the image streams on flat screens. The first step was to transform the panospheric image into an array of "constant angle". In this format, a delta-pixel change in the horizontal direction represented a constant change in angle. The same delta change in the vertical direction also corresponded to the same change in angle. For example, a 360 by 100 pixel image could be constructed to represent the full 360degree panorama with -90degree to +30degree elevation, with every pixel representing the view at a particular degree. As the actual camera provided 256 pixels over the 100vertical degrees, and from 1 to 1608 pixels over the 360degree panorama, the constant angle array created was 1080 x 300 (3x3 pixels each square degree). This transformation was accomplished with a look-up table that gave the destination x,y pixel coordinates for an i,j pixel in the raw image. Supersampling and subsampling was performed using linear interpolation and pixel averaging.

To display a segment of the imagery corresponding to a 60degree field of view, a 180 x 180 pixel subsection of the constant angle image centered around the desired direction could be displayed. However, the constant angle approach removes some of the normal distortion found in perspective projection images. While the center point stays in the center, the pixels moving outwards need to be adjusted from a linear to a spherical distribution. To solve the problem, a second look-up table was constructed that performed the task of adding back in the appropriate amount of distortion (the table was recalculated depending on the desired field of view).

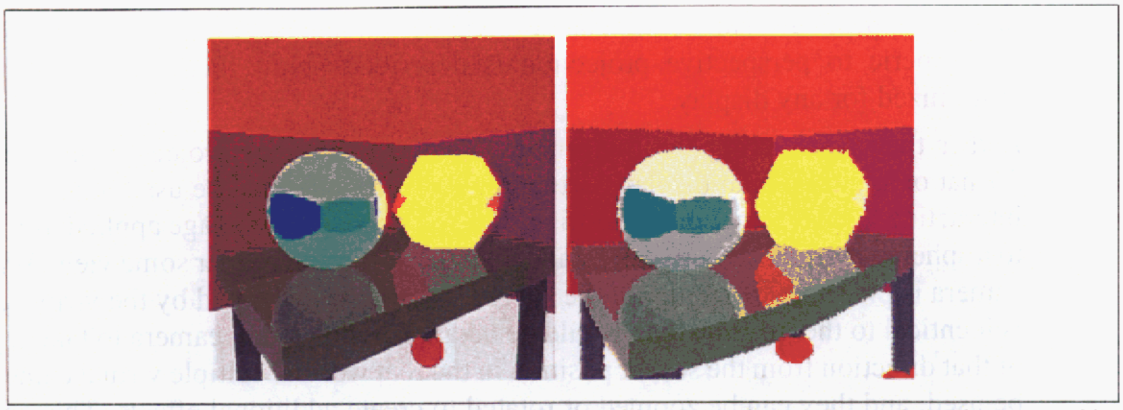


Figure 3-14: Comparison of perspective projection rendered image and constant angle region created by dewarping.

Figure 3-15: Dewarped strips from the prototype sensor, rendering environment, and final sensor version

The Telepresent Interface

Built on the virtual reality technique of a texture map applied to a a wireframe sphere, the Telepresent Interface was implemented to serve as the primary display method for the Atacama Desert Trek. Using Silicon Graphics hardware and Performer, a software application developed for 3D rendering, the Telepresent Interface was able to perform real time rendering of the panospheric video. The Telepresent Interface was coded by the University of Iowa GROK Lab, headed by Dr. Geb Thomas. The GROK lab also implemented VRML methods of displaying panospheric imagery that was not utilized during this research [Thomas98].

The Performer application provided the underlying mechanics and basic user interface features. After construction of the wireframe model, a bowl shape that represented the spherical field of view of the panospheric sensor (360degree horizontal from -90degree to +40degree), the panospheric image was applied as a texture. Ambient illumination was selected so that the texture map would be visible on the sphere regardless of the angle from which it was viewed. The wireframe could be rotated in two directions and translated along a single axis.

A virtual camera was attached to the environment at a fixed point along the translation axis. Backface culling was used so that the camera was always looking through the near side of the sphere and imaging what was on the inside of the furthest wall. The rotation and translation of the wireframe provided the apparent effect of pan/tilt/zoom. Because the panospheric image comes from a reflective surface, the image is a "mirror" image, and so was flipped across an axis to put the left view from the camera onto the left side of the wireframe. The texture map was applied to the model so that the "front view" of the sensor was visible when the wireframe was in its home position.

Implementing the Immersive Display

During the Atacama Desert Trek, the Pittsburgh based operations console was located in the Electric Horizon Theatre, developed by Spitz, Inc. The theatre consisted of a 32 seat bleacher arranged before a screen representing a subsection of a 30foot diameter sphere. Three projectors were used to provide a seamless display on the surface, and each seat was equipped with control buttons. The operating console was placed in front of the lowest row of bleacher.

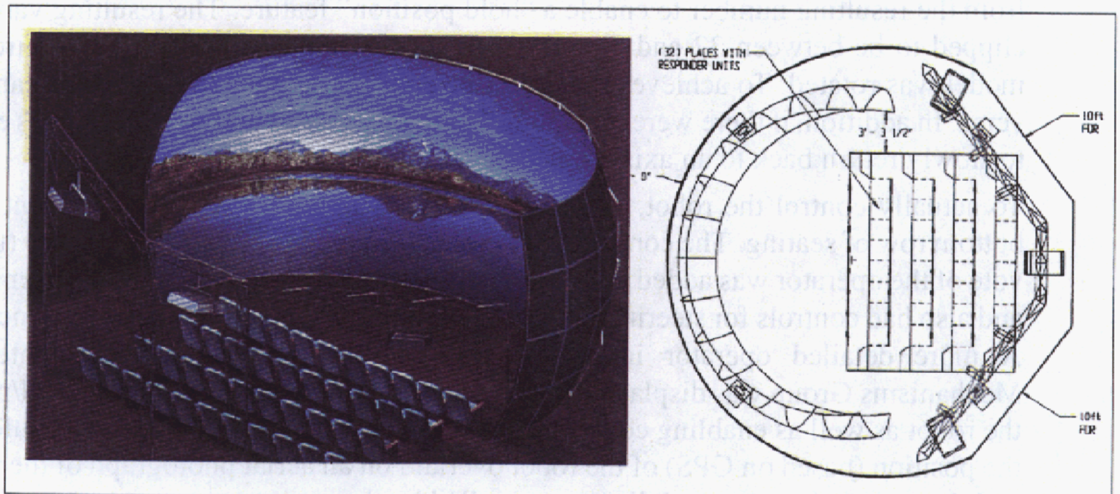


Figure 3-16: Artist rendering and schematic of Electric Horizon Theatre (courtesy of Spitz, Inc.)

The screen in the theatre duplicated a spherical segment comprised of a 200degree azimuth and a 60degree elevation. The screen was also tilted 15degrees downward, so that the field of view at the front of the screen was +15 to -45 and at the sides of the screen was +32 to -28. Three CRT projectors were mounted behind the bleachers. Each projector was responsible for covering the entire vertical field of view (60degrees) and over a third of the horizontal field of view (80degrees). The overlapping regions of the projectors were cross-faded in intensity so that the edges were seamlessly blended. Operating at a resolution of 800 x 600 pixels, the projectors provided a resolution of 2000 x 600 pixels over the entire screen.

The audience seating was a bleacher tier consisting of four rows. The center of the sphere created by the display screen was located at eye level of people seated in the top row, but directly overhead of the second row. The nominal seating capacity was 8 people per row, but 12 people could fit comfortably. Each row had eight button boxes distributed over the row. The button boxes were equipped with a red, a green, and a yellow button arranged horizontally. The SGI workstation that generated the signal for the display was capable of reading the status of each button individually.

An SGI Onyx computer with four R10000 processors and 64megabytes of texture memory was used to generate the display. The computer monitor was set at a resolution of 1600 x 1200 and divided into quadrants. The upper left, upper right, and lower right quadrants were respectively sent to the left, center, and right projectors.

The Telepresent Interface was used to display the panospheric video in the Electric Horizon theatre. The virtual camera was replaced with a three camera cluster. The fields of view of the three cameras were set so that when the cluster was collocated with the center of the wireframe model, the cameras acquired the same fields of view on the sphere as the projectors were responsible for displaying. The three camera views were then displayed using the Performer application on the appropriate quadrants of the computer monitor, and as a result on the spherical screen.

The control buttons were used to enable a votive panning feature. The total number of left buttons and right buttons were subtracted to determine the audiences desire to look either left or right. The number of central buttons being pushed were then subtracted from the resulting number to enable a "hold position" feature. The resulting value was clipped to be between 32 and 0, and the larger the number, the faster the wireframe model was rotated. To achieve a pan-left, the sphere was rotated to the right, and vice versa. In addition, if there were no buttons being pressed, an automatic vote was entered to slowly return back to an axis aligned or "front-facing" view.

To actually control the robot, a control console was built and placed in front of the bottom row of seating. The console had buttons for panning/tilting the imagery (the pan vote of the operator was added in as a vote of 16 with the button scheme when enabled), and also had controls for steering the robot left and right and issuing an emergency stop. A more detailed operator interface developed by the NASA Ames Intelligent Mechanisms Group was displayed on a computer monitor, and showed the roll/pitch of the robot as well as enabling control of the robot speed. This interface also could show the position (based on GPS) of the robot overlaid on an aerial photograph of the region, and the current location and distance travelled by the operator.

Chapter Summary

- A panospheric optic comprised of a hemispheric mirror is a suitable choice to generate the field of view desired for telexploration during the Atacama Desert Trek.

- Modeling, simulation, and prototyping of the optic were used to create and verify a working design, which was then refined and miniaturized.
- The complete panospheric sensor, consisting of optic, camera, packaging, and processing components was implemented using a combination of off-the-shelf and custom components.
- The distorted image acquired by a panospheric optic can be dewarped using point processing of the pixels, or through the use of virtual reality style processes, into whatever format is most suitable for the display medium.
- The Electric-Horizon theatre was the principal display medium for telexplorers during the Atacama Desert Trek, and the telepresent-interface was utilized to perform the displaying of the acquired video.

Chapter 4

Evaluating Panospheric Telepresence

The process of **Evaluating Panospheric Telepresence** involved choosing a set of display modes, observing and debriefing the operators of Nomad, and interpreting the data for results.

Seven different modes of displaying panospheric video in an immersive theatre were implemented. The video was displayed on four different environments, but only the immersive theatre was used as a basis for comparison of the display modes. Observations from the various displays were collected, along with comments and thoughts from the telexplorers.

Observations were principally focused on how well the telexplorers were able to make progress across the desert. Additionally, qualitative notes on the comfort level of each telexplorer were taken, based on physical actions, verbal comments, and duration of their use of a specific display mode. Observations showed that peripheral imagery and resolution affected teleoperation significantly.

The value of panospheric video was demonstrated, both for telexploration and for the more restricted task of remote science. During telexploration with access to panospheric video, novices and experienced operators had different expectations, but surprisingly similar abilities. The preferred mode of display was to provide peripheral information, but to place both front and peripheral video in the foveal viewing region of the telexplorer.

From the interpretation of the observations, the thesis assertion was deemed to be valid.

Modes of Operations and Display

Throughout the Atacama Desert Trek, telexplorers were observed driving the robot. Observation began during the initial process of training the telexplorer about the telepresent interface and operator console. The observation continued during actual telexploration, and ended with an opportunity for them to give comments and a debriefing concerning the experience. Six modes of display were available for each telexplorer, although most only operated the robot under one or two. Four of these views were originally intended, with the other two made available through the telepresent interface.

Observation Process

Unique to the Atacama Desert Trek was the primary use of novice operators to perform telexploration. This meant that most of those being observed had never previously operated a robot in any capacity, and had never been involved in a remote exploration. Even those who were experienced robot operators were somewhat novices, as the transforming locomotion system of Nomad was also unprecedented. There were three individuals who started as novices and became the principal expert operators of Nomad during the course of the Atacama Desert Trek.

There was no screening process for selecting telexplorers, as they were randomly chosen from those who expressed an interest. On average each telexplorer had control of the robot for five to fifteen minutes. "Training" was limited to a short explanation of the operation console, and the telexplorer was given a task statement dependent upon the current situation (drive in general compass direction, drive toward a landmark, extricate from trouble, etc.).

The controls for operation included specifying a turn vector, adjusting the throttle and an emergency stop. For safety reasons, the robot had an automatic time-out which would send the emergency stop command if there was no operator input for a 30 second duration. The telexplorer could also pan and tilt the imagery if the display mode permitted these actions.

One of the display modes described below was provided for the initial telexploration experience. As many of those who operated the robot spent time in the immersive theater prior to being selected, they were familiar with several of the display modes. After attempting operation in the provided display mode, the telexplorer was allowed to request one of the other display modes if they desired.

Both success in telexploration and comfort of the telexplorer were observed. Success refers to the ability of the operator to select a viable route across the unknown terrain and/or their ability to extricate from untraversable terrain conditions encountered either by themselves or as the result of the previous operator's efforts. Comfort was judged based on verbal comments, amount of time spent adjusting imagery, overall awareness for the robot's environment, and their ability to interact well with those just viewing the imagery. The average speed of the robot and the smoothness of the speed profile was also observed as a measure of both telexploration success and operator comfort.

At the conclusion of the telexploration experience, the operators were given the opportunity to make any comments, or ask any questions. Likewise, they were debriefed about their thoughts on the experience, and their satisfaction and comfort with the display modes available for their experience both in controlling the robot and viewing the video.

Display Modes

Operation was performed in six modes, ranging in immersive capabilities. When the research was proposed, there were four intended display modes. During the development of the telexploration experience two "advanced" modes were added to these basic modes. The basic modes allow comparisons between immersive panospheric imaging and the traditional teleoperation acquisition/display modes of a single fixed camera, a pan/tilt camera, and a camera array. The two advanced modes take advantage of the telepresent interface capabilities and replicate an overhead or "bird's eye" view, and an "off-board" or third-person point-of-view. Sketches of the six modes as seen on the electric horizon theatre screen are shown in Figure 4-1.

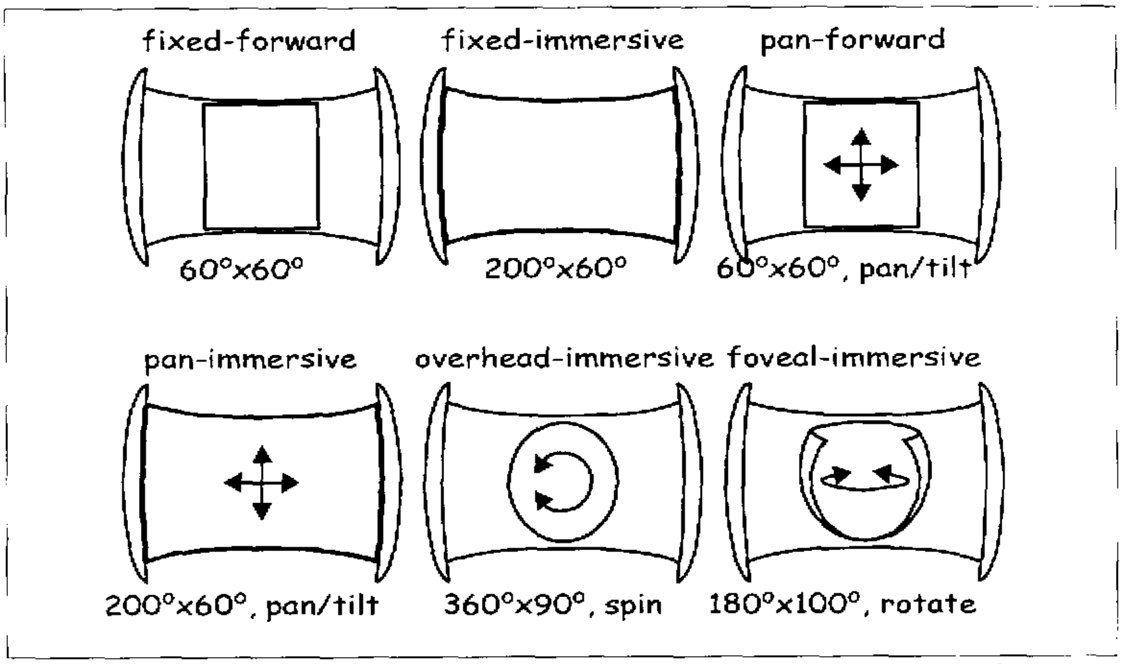


Figure 4-1: Sketches of the display modes used to observe and evaluate panospheric video for telexploration.

Basic modes

The simplest mode was the "**fixed-forward**". The view available to the operator was the forward looking view from the robot. For this mode, only the middle of the three projectors that comprised the Electric-Horizon theatre was kept active. This generated an 80degree horizontal by 60degree vertical field of view. Due to the edge-blending employed to make the total display mostly seamless the sides of the view faded to black rather than cutting off sharply, leaving a usable 60degree horizontal field of view. There was no ability to pan or tilt the video, and the view extended from the hood/wheels of the robot up to 20degrees above the horizon. The audience had no ability to interact with the video.

The "**pan-forward**" mode extended the fixed-forward by allowing the telexplorer to pan and tilt the video field of view in any direction about the robot. The configuration of the Electric-Horizon theatre was kept with the only the center projector active. The audience had the ability to concurrently pan/tilt the video using the votive button scheme, although the telexplorer's vote could override all but a near unanimous audience decision. Unlike a real pan/tilt camera, the commands to adjust the views were instantly provided, rather than having to wait through the time-delay for the robot to receive the command and actually servo the mechanism.

Immersive display modes were accomplished by adding in the two side projectors. The "**fixed-immersive**" mode enabled a full 200degree horizontal by 60degree vertical field of view. Neither the telexplorer nor the audience had the ability to adjust the displayed fields of view. This mode was essentially a mimic of an immersive view created by a forward looking camera array with seamless registration.

The last of the basic modes was the "**pan-immersive**". Using the same display hardware setup as the fixed-immersive, this mode presented the telexplorer and the audience with the same image manipulation capabilities as the pan-forward mode. The immersive video could be instantaneously panned and tilted to place any view vector at the front center of the theatre.

Advanced modes

The pan-immersive mode was originally conceived to be the primary evaluation for panospheric telexploration, but with the telepresent interface, advanced modes that further exploited the capabilities of the panospheric sensor were enabled.

The "**overhead-immersive**" mode most closely resembles a bird's eye view of the surroundings of the sensor. Rather than keeping the eye-point in the center of the texture-mapped sphere as in the four basic views, the eye-point was lifted straight up above the center and rotated straight downwards. The resulting view was the full 360-degrees around the sensor, and ranged in elevation from -90degrees (straight-down) at the center of the image, out to the horizon at the edge. This circular image was then displayed filling the center region of the Electric-Horizon theatre, covering what would be a 60degree by 60degree viewing region in the fixed or pan-immersive modes. The image could be panned, but not tilted, with the result of a pan being a spin of the image. The view to the front of the robot was always visible somewhere, but could be "upside-down" if the image was panned/spun around to position a back view at the top-center of the display.

The final display mode was the "**foveal-immersive**". Like the overhead-immersive mode, the virtual eye-point was pulled from the center of the sphere, but this time outwards instead of upwards. The resulting view showed the entire front hemisphere of the video (180degrees horizontal, by -90 to +40degrees vertical). Informally called the "cup view", the mode was similar to looking at a cross-section of a bowl with the video displayed upon the inside surface. Positioned at the center of the Electric-Horizon theatre display, the entire hemisphere was visible in a 80degree horizontal by 60degree vertical region. The mode again allowed for pan, but no tilt, with the effect of panning being an apparent rotation of the "cup".

An additional effect of these two advanced modes was that the apparent resolution of the display increased. Because the sensor did not have the resolution to match the display at the basic-immersive scale (roughly 540x120 pixels of data on the sensor were mapped to 2000x600 pixels on the display) the basic modes looked smoothed and blurry. Because the advanced modes both showed more field of view, and displayed it in a reduced portion of the display, the mapping became nearly direct, and thus the video appeared to have sharper resolution.

Details of the Observations

Observations are reported in general for each of the display modes, and then more specifically with respect to achieving progress during the telexploration and the effect on the operator comfort. Also included is a summary of observations from the NASA Ames telescience experimentation week. Figure 4-2 provides the observation details in table format..

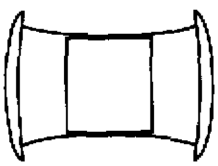
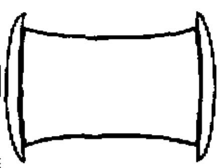
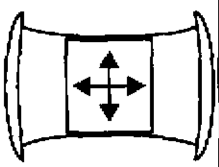
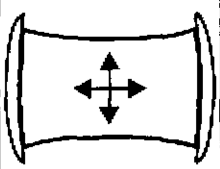
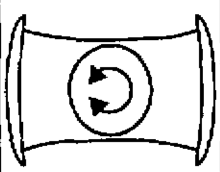
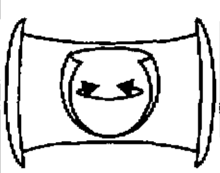
mode	behavior	awareness	metrics	reactions
	turning in circles cautious operations frustrated operators	minimal awareness as expected	stops: 5-10sec restarts: 20-??sec speed: 20-30% distance: ~20m	average avoidance horrid extrication
	smooth arc paths cautious operations intimidated operators	good awareness except for finding a feature off to the side	stops: 50-60sec restarts: 20-30sec speed: 20-30% distance: ~50m	good avoidance tedious extrication
	straight line paths reckless or cautious inattentive or deliberate operators	confusion about features being on the left or right	stops: 30 or 10sec restarts:90 or 20sec speed: 90 or 30% distance: 50 or 25m	average avoidance average extrication
	spiraling paths cocky operations confused operators	no real awareness except for extrication too much panning	stops: 10-30sec restarts: 5-30sec speed: 70-80% distance: ~90m	horrid avoidance average extrication
	smooth arc paths with frequent reverse bumper-car style bewildered operators	ravine depths not easily distinguished objects easily seen	stops: 40-60sec restarts: 10-30sec speed: 80-90% distance: ~140m	bad avoidance (of ravines) good extrication
	graceful paths effective operations excited and involved operators	fantastic awareness	stops: 60-90sec restarts: 5-20sec speed: 85-95% distance: ~160m	excellent avoidance excellent extrication

Figure 4-2: Table summarizing the observations

Synopsis of Each Display Mode

Fixed-forward: Serving as the baseline display mode, expectations for performance were low, and indeed, the performance was poor. Telexplorers consistently had trouble understanding the environment, and spent a lot of time slowly turning the vehicle in circles to try to get a sense for what was around the robot. Most of the telexplorers felt very nervous and concerned about the robot during this mode, and several admitted to being scared they were going to crash. Hitting the stop button was very common, as was getting stuck on slopes and soft soil. The path generated by telexplorers using this mode was almost like "connect-the-dots", as the robot would be sent in some direction, and then would circle until the telexplorer was comfortable with the new path, and take off again.

Pan-forward: Expectations were slightly higher for this pan/tilt imitation, but the results were still rather poor. Telexploration styles tended to fall into one of two categories. The first group would drive rather recklessly, panning the imagery around, while driving nearly full speed. They often got stuck in soft soil and spent a significant amount of time (up to a minute) before they noticed the vehicle was not moving. They also often would leave the vehicle sitting still as a result of the time-out feature while they panned about. The second group was much more cautious, preferring a "step and look" style. After commanding the vehicle to move, they would stop it after about 15 seconds and sit still while panning about trying to decide where to go next. Then they would issue a command and repeat the process. This generated a path that was like creating free-form polygons, as it did not have the circles of the fixed-forward mode, but still had the fairly obvious "steps".

Fixed-immersive: Telexplorers fared pretty well in this mode, but were rather vocal about their dissatisfaction with it. Common complaints were with the resolution being too blurry (although it was the same as in the two previous modes and there were few comments about resolution in those modes), and the display being intimidating with video all around. Operation itself was fairly smooth, and driving in arcs was more the style with telexplorers making corrections during operation rather than stopping and adjusting. There were fewer incidences of getting into trouble than with either of the two previous modes, with misunderstanding the time-delay the main problem of the telexplorers.

Pan-Immersive: This mode was anticipated to produce the best results of the four basic modes, but in fact, was not as effective as the fixed-immersive. There were the same complaints about the display being blurry and intimidating, but the main problem was the telexplorer's tendency to confuse panning with steering. And also, while the telexplorers said they felt very in control and very aware of the environment, they often ran into problems due to this cockiness, by not paying enough attention to the slopes or the terrain in the direction the robot was headed. The one clear area where this mode succeeded was in extrication, where very little time was needed to figure out what direction to go to get back to a safe area. However, correctly mapping that direction to a vehicle steering command was occasionally troublesome.

Overhead-Immersive: The first of the advanced modes, the results were mixed. The main problem was with navigation in sloped areas, as telexplorers were unable to get a sense for the scale of steps or ledges. In several cases, operators were unable to tell that the vehicle was stuck in one place, spinning its wheels into soft soil. Other than that, operation was very smooth, even when operating at top speeds. Telexplorers were able to make turns without slowing down and had a good sense for the direction they needed to go to reach a landmark, unless the landmark was obscured by nearby terrain at which point they would get lost. Telexplorers were also very quick to notice anything going on in any direction, such as support personnel coming into sight. Occasionally this resulted in quick redirection of the robot to chase these personnel, and rather extended games of "bull-fighting" ensued upon occasion.

Foveal-Immersive: Unquestionably, the best of all the display modes. Telexplorers had no complaints about the mode, and it was almost always the mode requested by a telexplorer wanting to "drive the robot the right way". Avoidance of bad terrain was the best of any mode, and extrication was also very quick. There was great understanding of slopes and ability to determine if upcoming ditches were too deep to traverse or not. 90% of the traversal progress of the Atacama Desert Trek was made by telexplorers using this display mode. There was less confusion between panning and steering than in the pan-immersive mode. Operations were in general performed at top speeds, and telexplorers seemed to be able to adjust to the time delay without as much problem.

Traversal Progress

Selecting a viable route

When the task given the telexplorer was to navigate to a landmark or along a general compass direction, they were responsible for finding a route that would allow safe passage. In the Atacama Desert, the principal barriers to progress were slopes of soft soil that Nomad was unable to climb, and ravines formed from ice melt-off that could be over fifteen feet deep. Nomad was limited to 30inch drops and 15degree slopes. While there were a few telexplorers whose primary interest was to try and find ways to break the safeguarding system, the majority of those who operated the robot were actively trying to avoid hazardous and untraversable terrain.

Expectations were that the immersive display modes would allow higher success at avoiding trouble. Thus, it was not surprising that the fixed-immersive and foveal-immersive modes did have the best successes in anticipating trouble. Telexplorers did occasionally have difficulty predicting how soon to turn due to an unfamiliarity with the time-delay. However, most telexplorers initiated a sequence of turns when they first took control of the robot and thus knew to slow down or stop when the terrain transitions from flat (and completely safe) to rugged. There were very few collisions when using the foveal-immersive mode or the fixed-immersive mode.

The overhead-immersive mode suffered from a lack of ability to discern slopes and a lack of ability to see above the horizon. Telexplorers commonly ran the vehicle straight at dunes and had to put the vehicle into reverse after getting it stuck in the soft sand on the slope. Also, they tended to lose track of landmarks when they got eclipsed by dunes and occasionally got very lost or ended up heading in an erroneous direction.

The pan-immersive mode should have been as good as the fixed-immersive mode, but the added ability of the telexplorer to look around them acted as a distraction rather than a benefit. It was very common for the telexplorer to become so used to panning the imagery from side to side that when it was necessary to turn the robot they would first

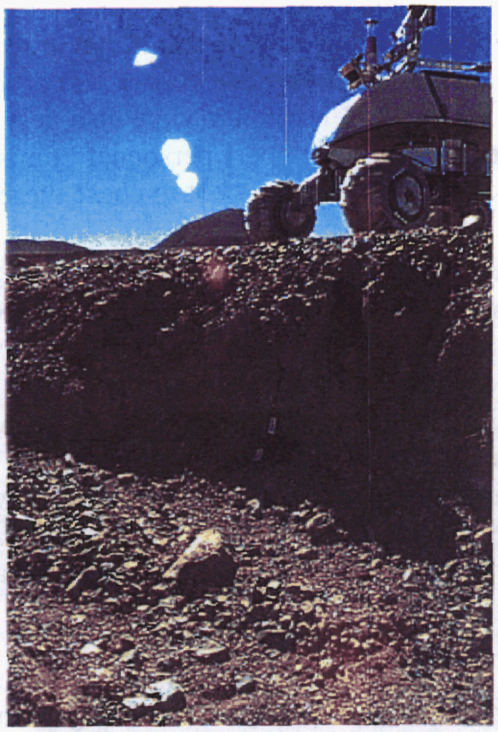


Figure 4-3: Nomad perched atop the edge one of the many dry washes encountered during the Atacama Desert Trek telexploration

turn the imagery in that direction as though expecting the robot to follow. The telexplorers using this mode had the worst time becoming accustomed to the time delay with steering, as the panning was instantaneous and the steering was not. The result was almost a form of bumper-car driving, where the robot would bounce from one difficulty to the next in a most ungraceful fashion.

The two limited field of view display modes, fixed-forward and pan-forward, did not lead to as many problems as either of the overhead-immersive or pan-immersive modes, but in the case of the fixed-forward mode, that was more due to an inability to get the robot to go anywhere rather than a success at avoiding trouble. The fixed-forward mode required telexplorers to turn the robot to see what was ahead to the left and right, and often times resulted in steering directly into trouble in an attempt to see what was nearby. The lack of a point-turn capability of the locomotion system acerbated this situation.

The pan-forward mode was perhaps the third best of the display modes at avoiding trouble and in deciding which direction to go. The step and look strategy that resulted for almost all telexplorers is a common approach in normal teleoperation with time-delay, and has been used with successful, albeit slow, results in many environments. The limited field of view prevented the telexplorers from ever feeling very confident about the environment, and so operations were rarely at full speed when moving, and over 80% of the operation time was spent with the robot stationary. There was a small group of operators who decided to trust the safeguarding system and just point the robot some direction and let it run free while they panned around to see what was interesting.

This mode led to frequent troubles, since by default, the robot would stop whenever the safeguarding system detected a problem, or when the vehicle would get stuck in soft soil that the safeguarding could not detect.

Extrication from untraversable terrain

In an exploration setting, it is almost inevitable that a promising path will turn into a box-canyon, or that what appears to be a hard slope is in reality a soft dune. Thus, extrication from trouble is vital for telexploration success. Often, extrication requires seeing what is to the sides and back of the robot, and thus the expectation was that display modes providing a backwards view would be more successful.

As expected, fixed views were totally unsatisfactory. The fixed-forward mode was incredibly useless, and extrication required switching to a different view nearly 95% of the time. The immersive-forward mode was indeed better, and telexplorers were able to extricate from trouble without switching views, although several attempts were required on average. Often times, simply driving straight backwards would have been a good way to get out of trouble, and telexplorers opted to try to turn the vehicle around and got into a worse situation. When driving blind, telexplorers were much more conscious of the tilt and roll indicators, and often times watched only those rather than utilizing any imagery.

Also as expected, the pan-forward mode was effective, although requiring significant amounts of panning. The advantage of the panospheric sensor was apparent in that the telexplorer could perform the pan instantaneously rather than having to wait for the command to reach the robot and be interpreted and the new video returned. This allowed for fairly quick extrications, and telexplorers tended to avoid driving into worse trouble while trying to figure out which way was safe to go.

The three most immersive display modes had the expected success in extrication ease. The pan-immersive was satisfactory, and really was not that much different than the pan-forward display mode, as most operators tended to pan around until the desired direction was centered on the display. The main distinction between the pan-immersive mode and the pan-forward mode was that the telexplorer often planned a route out of the situation before issuing commands, rather than merely picking a direction to head. This frequently resulted in fewer steps to achieve extrication, but generally took longer because the telexplorer would spend more time making the decision. Had the time delay been on the order of minutes instead of seconds, this distinction could have been much more of an advantage.

The overhead-immersive and foveal-immersive modes were quite different in achieving extrication. In both modes, operators were very confident about what they should be doing. However, the overhead-immersive occasionally had difficulties in making progress along ravines without running up onto the sides and getting stuck again. This really did not seem to bother the telexplorers, as backing up and trying again with a slightly different direction was done quickly. Nonetheless, the bulk of "damage" incurred by the robot's locomotion mechanism occurred as a result of telexplorers bouncing the robot back and forth on ravine walls while using the overhead-immersive mode.

The foveal-immersive mode had no such problems. And while the mode did require some panning to get to the back views, the heightened apparent resolution of the video made telexplorers very confident on their decisions and interpretation of the surrounding environment and extrication was quickly accomplished. Like the pan-immersive mode, operators tended to have a longer range plan for extrication with the foveal-immersive view, but were not afraid to start driving in the proper direction and make the final decisions as more information became available.

Average operating speed and traversal path

The last of the observations relating to achieving progress during operations focused on the operating speed and the characteristics of the path blazed by the telexplorer. The expectation was that speed would be determined mostly by the terrain, rather than by the display mode. It was also expected that the more immersive the display mode, the smoother the path would be. While the second expectation generally held, the operating speed was actually more dependent upon the display modes with telexplorers clipping along at full-speed for half of the display modes, and keeping the vehicle much slower for the other three. Taking into account the time spent trying to extricate, the average progress speed was very much dependent upon the display modes.

The foveal-immersive mode was nearly always taken to full-speed, and there were very few incidences of trouble that resulted in a slowing of speed. Likewise, telexplorers would drive at full speed when using the overhead-immersive mode. In this mode, however, there were frequent occasions of getting stuck on slopes, or careening into the walls of ravines. This led to a need to slow down or stop, back up, and continue again. Even after two or three troubles, the telexplorers rarely operated at a slower speed, and this was the only mode where telexplorers were urged to "take it a bit slower" through areas of rough terrain.

The pan-immersive mode was the other display mode operated consistently at high speeds. The telexplorers overall progress was similar to the overhead-immersive mode, as it was common to drive into trouble areas, and need to spend time performing extrication. Unlike the overhead-immersive mode, operations here would converge to a slower speed in rough terrain without external prodding.

The other three modes were operated much more conservatively, with speeds normally averaging about half to two-thirds of maximum. The only exception was the occasional telexplorer in the pan-forward mode that liked to just let Nomad drive itself while panning about oblivious to where the robot had been sent. Ignoring this subset of telexplorers, when taking into account average progress speed, the pan-forward mode was just as effective as the either the overhead-immersive or pan-immersive modes, since there were fewer incidences of trouble requiring extrication maneuvers, and extrication itself was pretty quick. The fixed-forward and fixed-immersive modes were slow in speed, and slower in progress due to an inability to extricate or to even decide which direction to travel next without staying in the same location for several minutes.

The path described by the telexploration had two basic forms: connected lines or smooth arcs. The fixed-forward, pan-forward, and pan-immersive modes tended to leave more of a "drive straight and then turn" path, while the fixed-immersive, overhead-immersive and foveal-immersive modes exhibited a smoothed route. The fixed-forward mode was distinct in that it was marked with "donuts" at almost every turn point as the telexplorer circled the robot to see which way to go next. Likewise, the overhead-immersive commonly left a path like a "jagged lightning bolt" with regularly spaced moments of backing up to get off of slopes and then trying again. The pan-immersive mode also produced an interesting path, as telexplorers seemed to have trouble figuring out the direction to take during extrication, and often generated "stars" by choosing a direction to retreat, realizing it was wrong, stopping, driving forward but again confusing left with right, etc.

Operator Comfort

In addition to the technical observations of how well telexplorers were able to make progress across the unknown environment, how the modes affected the telexplorers themselves were qualitatively observed.

Understanding of the time delay

There were three features of the Atacama Desert Trek and Nomad itself that made operations difficult. The first was the time delay, as commands to the robot were generally implemented in just about a second, but receiving, processing, and displaying the video showing that implementation took just over two seconds. Second, the transforming locomotion chassis was unusual in that the vehicle needed significant time to obtain the desired steering angle (up to ten seconds depending on the severity of the turn), and just as much time to return to driving straight. The third was that the telexplorer received no true indication of slope until the vehicle was actually climbing or descending it.

None of the basic display modes really helped much with the time delay, but the overhead-immersive and foveal-immersive did have an advantage. With more than a 60degree vertical field of view shown always, telexplorers could see the front wheels adjusting to the turns and straightening out. Telexplorers also had the field of view to get a good sense of how fast they were turning. After a short time operating in either of these modes, telexplorers began to send commands to the robot that correctly anticipated the time-delay.

The fact that it took so long for the robot to actually turn resulted in rather consistent oversteering at first. Telexplorers in the pan-forward or pan-immersive modes tended also to try to pan over to where they wanted to be heading, and wait for the robot to catch up. This usually resulted in the robot being driven in a U-turn, circle or figure-eight because the telexplorer eventually got confused about whether they were panning or whether they were steering. Many even forgot to ever straighten out of a turn and would spend two minutes just circling about, while panning and trying to adjust the speed so that the robot would actually go somewhere. In the fixed modes, getting lost in the video was not an issue, but telexplorers never really were able to figure out how long the locomotion chassis was taking to accomplish the desired turn, as the wheels were only marginally visible in the vertical field of view.

Awareness of the environment

To get a sense for the situational awareness of the telexplorers, they were asked questions to gauge their understanding the objects in the near surroundings, and understanding the scale of those objects. The primary features of the landscape included color changes in the soil, gradually sloping sand dunes, and the ravines eroded by ice melt-off. The biggest danger to the robot were the ravines with their sharp drops and corresponding high side-slopes.

Most telexplorers were able to understand the tilt/roll concept pretty easily, although there was often confusion about whether the vehicle was tilted forward or backward. There were multiple occasions during the trek when the safeguarding system failed to detect a fairly deep ravine, and the robot was driven over a substantial drop (in excess of three feet). Most of the time, the telexplorers were unaware that a big drop was coming, as they were not sure on the scale of the desert and somewhat trusted the safeguarding system. The foveal-immersive and the fixed-immersive were the only two modes in which telexplorers consistently noticed ravines and took successful action to avoid them (fixed-forward had too limiting a field of view to see the ravine, the pan-modes often had the telexplorer looking the wrong way at the crucial moment, the overhead-immersive indicated the ravine but not its depth).

The fixed modes were expected to have worse awareness than the ones allowing pan, however the pan-forward and pan-immersive modes actually generated the worst awareness. The fixed modes required the telexplorer to pay more attention and they were usually able to talk about the terrain they had driven over, or correctly predict what was going to be encountered by the next telexplorer. The basic pan modes acted more as a distraction, and the telexplorer would remember having seen a rock or a mountain, but could not recall if it was on the left or right of the robot. The telexplorers using the basic pan modes often were unable to even say which path they had taken.

The overhead-immersive and foveal-immersive views were great at seeing what was around the robot, and with the exception of interpreting slopes, the overhead-immersive was the best at accomplishing near-robot awareness. Telexplorers using the overhead-immersive mode were repeatedly able to return the robot to a particular location to get another view of a rock feature, or to successfully spend considerable time bull-fighting with support personnel who were accompanying the robot, and trying to stay out of the way. The foveal-immersive view also gave people a good sense of what was nearby, but was notable more for its ability to provide long-range awareness of landmarks or mountains, etc.

Frequency of panning

Most of the time the telexplorer took control of the robot, their tasks could be accomplished with a minimal of panning, regardless of the display mode. In the fixed modes, panning the video required actually turning the robot. Due to general curiosity, and often because they could, telexplorers always took time to pan the video.

In the fixed-forward and fixed-immersive modes, telexplorers generally gave up on the panning concept after driving in a couple of circles. However, after that experience, they generally had lost track of their original heading, or could have gotten into trouble. As a result, the telexplorers spent probably about half to a third of their operating time turning the robot for the purposes of getting more information about the surroundings. In the case of the fixed-immersive, telexplorers were able to get back on track fairly easily and tended to settle into a pattern of a slalom or sinusoidal or straight-loop-straight-loop path. The fixed-forward mode usually degenerated into loop-loop-loop-

loop with occasional attempts to get out of the loop and head in the right direction resulting in oversteer and the need for more loops to figure out what went wrong.

The pan-forward mode was interesting in that most telexplorers were reticent to pan away from the front view for long periods of time. They would take a quick peek to the left or right, but unless they were in trouble would pretty much come right back to the front. Most of the time, they would stop the vehicle before panning about. The pan-immersive mode led to significantly more panning, and the majority of it done when the vehicle was moving. The front view was usually still visible somewhere on the screen, but the telexplorer's attention was not often directed that way.

The overhead-immersive and foveal-immersive modes were almost always spinning about. There was never any loss of data during panning the overhead-immersive mode, and so telexplorers felt very safe in manipulating the view while operating. Interesting enough, it was pretty hard to tell if the robot was turning while panning either of the modes, but telexplorers rarely left the robot circling around and spent most of their time driving in the right direction.

Sharing well with the viewers

In addition to the telexplorer who was the only one able to control the robot, several of the modes allowed the audience in the theatre to participate in manipulating the video. In the pan-forward, pan-immersive, overhead-immersive, and foveal-immersive the votive buttons were occasionally enabled for audience pan (but with no tilt ability). This presented some interesting conflicts when the audience desired to view something different than the operator. While the operator did have a strong enough vote to dominate the decision, there was still some friction at times.

In the fixed modes, the audience tended to pressure the operator to drive a certain direction, or yell out that they were going in circles, or heading into a ditch, etc. This did at times help the telexplorer, but occasionally would result in the telexplorer giving up on operating the robot and spend more attention on the hecklers.

The overhead-immersive was the least affected, as the audience could never eliminate the view being used by the telexplorer, although they could rotate it so that it was upside-down or otherwise harder to see.

In the pan-forward mode, the audience was generally considerate of the telexplorer's requests to let them see what was going on while the vehicle was moving, and once the vehicle stopped, the audience and the telexplorer often worked as a team to decide where to head next. In the pan-immersive mode, the audience figured the telexplorer had a lot of video available, and would consistently force the "straight-ahead" view to the sides of the screen. The audience also rarely remembered that the operator needed to see the area the robot was headed into when it was turning, and as long as the front of the robot was still on the screen, did not feel as though they were impeding progress.

The audience enjoyed to pan the foveal-immersive mode when the robot was not moving, but were willing to sit back and be more of a passive audience during actual operations. In general, this view was more attractive than any other, and many people commented that they could sit and watch it for a long time without feeling bored because it was a view they had never experienced before.

NASA Ames Telescience evaluation

The NASA Ames Intelligent Mechanisms Group participated in numerous aspects of the Atacama Desert Trek. One of the main focuses of their participation was in the performance and evaluation of telescience. The timing of the Atacama Desert Trek enabled researchers to prepare for the Mars Pathfinder landing and provide some form of baseline for telescience accuracy, since geologists could go to Nomad and record actual ground-truth to compare to the telescience findings.

To assist with the remote geology experiments, a pan-tilt head was fitted onto Nomad that carried two pairs of cameras. A narrow field of view monochrome camera pair (35degrees) served in the traditional role of a pan-tilt camera for telescience (but providing stereo imaging to allow terrain reconstruction and range calculations). In addition, an extremely narrow field of view camera pair (11degrees) was used in an attempt to directly reproduce human foveal vision, having 0.30milliradians per pixel.

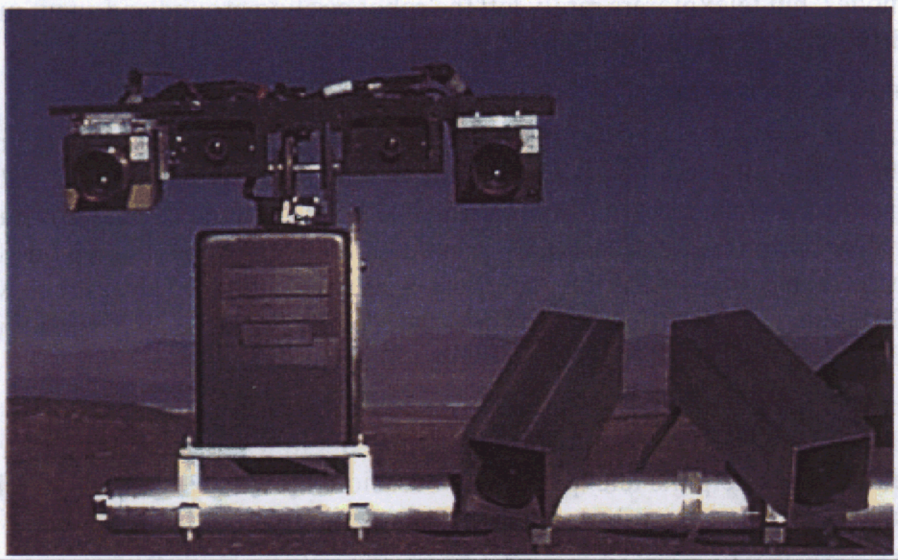


Figure 4-4: The pan/tilt cameras utilized for the telescience experiments. The cameras to the lower right are part of the stereo imaging system used for autonomous navigation and safeguarding.

The basic approach to acquiring awareness of the environment was to use the normal field of view camera pair to take a series of pictures and generate a full panorama about the robot. Then scientists would evaluate the panorama for where to go, and what looked interesting, and then the rover would be commanded to visit that area, and once there another panorama would be taken at a closer range to the robot. This panorama would then be used to decide where to direct the high-resolution camera pair for acquiring a precise look at the selected feature. Then the process would be repeated.

Experiments on panospheric video for telescience

There were several experiments performed that compared the effect of panospheric video on telescience, which have been documented by D. Wettergreen, et al. and which will likely be further detailed by Dr. N. Cabrol, the science lead for the experimentation. The experiments were divided into two principal operating scenarios: lunar and martian. In the lunar scenario, the operations were similar to that of the bulk of the Atacama Desert Trek. However, for the martian scenario, the time delay was increased to the fifteen minutes required for round trip transmission to Mars. This placed a premium on making correct decisions, and in minimizing the amount of time spent operating the robot or waiting for imagery to be acquired.

In the lunar experiments, the attempt was to perform "geology-on-the-fly", in which scientists were charged to assess traverseability and survey the basic geology while keeping the rover moving three-fourths of the time. This was designed to simulate a long-distance traverse between sites of suspected scientific interest. During this experiment, the Ames team established a record for remote exploration with the traversal of 1300 meters and the investigation of ten science sites [Wettergreen97].

In the martian test, the principal endeavor was to evaluate the usefulness of panospheric video when operating with time delay. Both with and without panospheric imagery, a science site was approached, enough imagery was taken to allow for an understanding of the geology and other terrain features, and traversal to another science site was performed. While using the panospheric image stream, the Ames team acquired fewer uninformative images, and twice as much area was examined in the same time.

Notes on the Display Media

Although the telexplorers observed during the research were limited to using the Electric-Horizon immersive theatre, two other display styles were utilized for the panospheric video during the Atacama Desert Trek. A flat-screen computer monitor served to preview the video before it was sent to the immersive display. The video was also broadcast for television viewing during RoverTV (also, the NASA Ames display was in effect a big-screen television).

Immersive Theatre

When displaying the panospheric video for the Electric-Horizon theatre, there was significant group interaction from the audience, the resolution and color were displeasing, the imagery was captivating and provided a strong sense of being at a remote location and yet immersion was not necessarily required.

The 32 member audience and the lone telexplorer who had control of the robot almost always became a team during operations. The telexplorer would often utilize the audience to perform image panning while concentrating on sending commands to the robot and watching the roll/tilt meters. When the telexplorer tried to take complete control, the audience generally grew restless and began to interfere in the process by panning the video in generally undesirable ways. The main exception to this was in the foveal-immersive view where the audience always kept a long attention span, even when comprised primarily of young children.

The resolution of the imagery was not sufficient to mesh with the resolution capable of the display, and as a result the full immersive views looked much worse than the foveal-immersive or overhead-immersive views. The colorization on the sensor itself was not well calibrated (no white balance) and so the video always had a tendency to look reddish. This enhanced the "remoteness" of the environment by twinging the sky a little away from blue and more toward purple. In addition, color-balancing the video between the three projectors was difficult, and a rather noticeable color difference was evident between the center and right projectors.

The initial question asked to every new audience in the theatre were "Can you tell me what you are looking at?" The most common answer after looking at the vegetativeless, barren landscape was: "the moon" or "mars". Even after learning that it was a South American desert, the audience was still excited by the concept of being a "telenaut". When the imagery was pulled back from the fixed-immersive mode to reveal the foveal-immersive mode, the audience was quite captivated. Some visitors to Carnegie Science Center stayed in the theatre for several hours, fascinated by the experience.

Computer monitor

The attending stateside operations team had access to a computer monitor on which the video destined for the immersive theatre was previewed. The operation console and the votive buttons, along with the multiple projectors were connected to this computer. While the monitor did not show anything more than the immersive theatre, it did display the imagery over a smaller field of view, and so the resolution had the appearance of being much sharper. The color of the monitor was also better than that of the projection system, and so subtle features in the terrain were easier to distinguish, as were objects in shadow.

Occasionally, interested visitors would be shown the computer monitor in answer to questions. After watching the monitor for a while, they would usually question whether or not the imagery was being generated by the computer as opposed to actually being taken live. This question of belief in the remote experience never happened with the immersive theatre views, but was much more common when confronted with what appeared to be a virtual-reality display on a computer.

Television and Big Screen

The most unexpected success of the Atacama Experience was with the development and production of RoverTV. A television broadcast was created out of the foveal-immersive and overhead-immersive display modes. These two modes were already known to be successful from their use in the Electric-Horizon theatre, but their utility in conveying the immersive experience on a flat screen was unanticipated.

The NASA Ames team was also without an immersive theatre for viewing the panospheric video, and utilized the same two modes on a large projection tv during their telescience experiments. The shape of the imagery displayed through the foveal-immersive mode translated very well as a three dimensional bowl on the flat screen, and operators were very easily able to tell that they were looking at both forward and peripheral image regions.

Interpretation of Results

At the highest level, panospheric video was highly successful in enhancing telexploration. Video conveying the entire surroundings of the robot had the anticipated benefits of improved awareness and extrication from untraversable terrain. However, there were unanticipated preferences, as well as some novel interactions between groups viewing the video at distant sites.

Panospheric vs. Traditional

The four basic display modes were designed to allow direct comparison between panospheric video and traditional imaging solutions for teleoperated robots. However, it was the two advanced views that provided capabilities beyond the normal teleoperated visualization techniques that really demonstrated the value of immersive video for telexploration.

As expected, teleoperation with fixed views exhibited weaknesses in route selection and extrication from trouble, even when possessing 200degrees of azimuthal field of view. Likewise, pan/tilt views resulted in improved extrication, but generally yielded poor progress. Allowing a vantage point with a bit more distance from the scene and including much more of the robot and the terrain (foveal-immersive and overhead-immersive) was vastly superior to the basic approaches.

The biggest advantage of the panospheric video was the instantaneous panning ability. Rather than having to suffer through the time delay for a change of view, the presence of the entire image meant that changing views required only processing time. With continual updates of frames, the view in every direction was kept current. In this regard there was never any need to work from a previous location's panorama or to remember what was to the left of the robot. This was a fantastic benefit to the Atacama Desert Trek, as it allowed novice drivers who were quite young (7 to 14) that normally have rather short attention spans to operate the robot successfully, as they could just check back to see what they missed at any time.

Preferred Display Mode

Without question the two preferred display modes were the advanced foveal-immersive and overhead-immersive. The principal reason given for the preference was one of improved apparent resolution, but in general the two modes were a bit more entertainment in style. The spinning cup view and the bird's eye view are not the normal views that humans have in everyday life, although they are views that are commonly used in arcade driving games and on maps.

While the overhead-immersive mode was used successfully to keep track of objects in the near surroundings, it was fairly common for the telexplorer to be tilting their head way to the side (and even somewhat upside-down at times) to keep the video from the front of the robot in the right orientation. The foveal-immersive mode always kept the standard orientation of the surroundings, and this was much preferred, even though it meant losing out on the back hemisphere, and occasionally not being able to see in front of the robot at all.

Novice vs. Experienced Operators

With the designed safeguarding system aboard Nomad, it was possible for novice operators to take control of the robot. In addition, there were several experienced operators who were familiar with previous mobile robot operations. As it turned out, the two groups had similar capabilities and results, but had very different expectations and impressions on the immersive video.

Both novices and experts had a slight difficulty adjusting to the time delay, although the experienced operators knew that they needed to anticipate commands. The principal obstacle was the unusual chassis which took a significant amount of time to fully understand. Very few novices were able to fully achieve the motion they intended, and experienced operators usually took over an hour to become proficient.

Where the two groups differed mostly was in expectation and impression of the video. Novices generally compared the experience to movies or television or computer games and found the resolution low, the color fairly dim, and the speed of the robot sloth-like. The experienced operators recognized that the resolution was indeed a bit low, but the field of view was substantially greater than in previous robotic systems, and were able to accept the trade-off. The walking-pace speed of the robot was not viewed as a detriment, and some experts remarked that slowing the system down further would

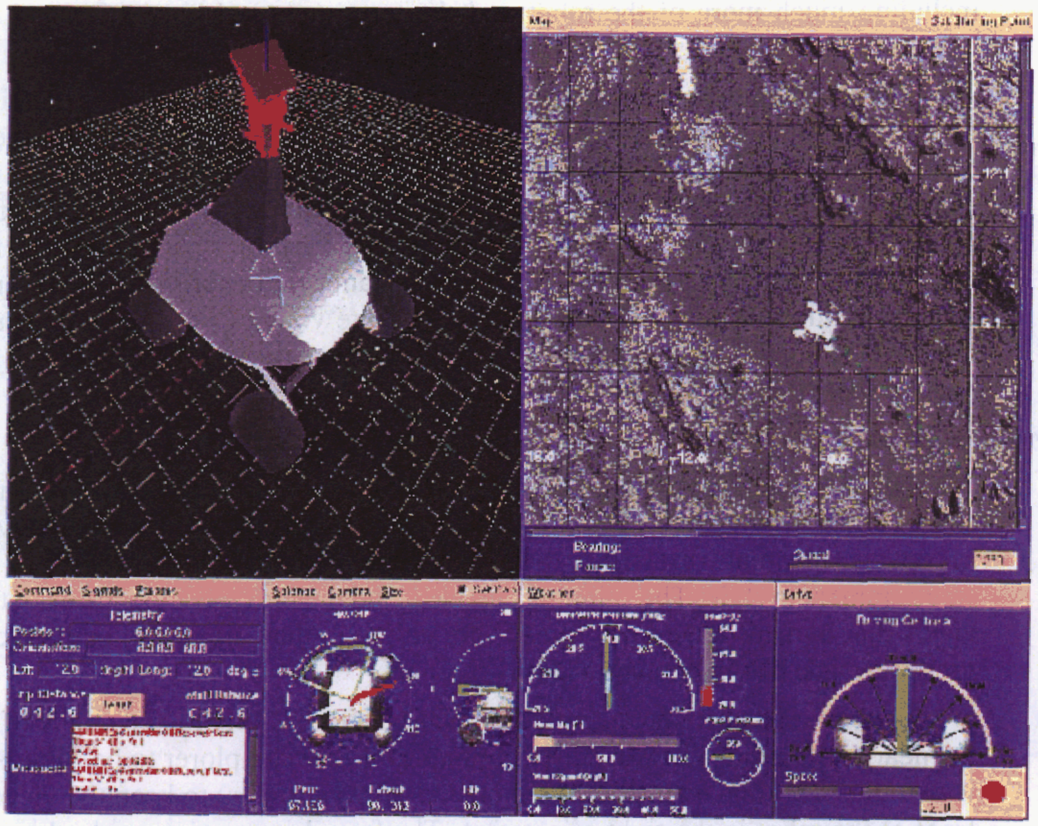


Figure 4-5: The virtual dashboard interface developed by NASA Ames IMG and used for controlling Nomad during the Atacama Desert Trek.

result in fewer instances of trouble.

Experience operators were less willing to trust the safeguarding system, and preferred to take the cautious approach, whereas the novices were more willing to plunge ahead with the assumption that the robot would do the right thing. In general, the safeguarding system was not fully able to keep up with the novice operators, and was more of a hindrance to the experienced operators.

Independent Simultaneous Viewing

Unique to immersive video and innovated for robotic telexploration on the Atacama Desert Trek, was the ability to independently select the observed field of view in independent locations simultaneously. During the Ames telescience experimentation, visitors to Carnegie Science Center were able to interact with the robot and control the pan and tilt of the imagery, even though the robot was being controlled half a continent away. Likewise, the imagery was viewed and manipulated and archived at Ames Research Center throughout the entire time that telexploration was occurring in the Electric-Horizon theatre.

One incident illustrates the value of independent simultaneous viewing. While operations were occurring in Pittsburgh, one telexplorer drove near to a small crater. Although the crater had failed to pique the interest of the audience in the theatre, it was considered a possible meteorite impact site by the Science team at Ames. They recorded the position of the robot at the time, and a short while later when they were given control, were able to head back to that location. While the crater had no significant value to one audience, it was rated highly by an independent audience with different training and interests.

NASA Ames conclusions

The Ames team had several conclusions from their experiments using panospheric video for telescience. In previous work with remote exploration, the team had documented the importance of panoramic imagery for robot localization, environmental understanding, and robot traversal. The panospheric image streams provided an elegant solution to this need, and allowed scientists to obtain a "sense of position and orientation at the remote site that has been lacking in previous field experiments and substantially benefits and accelerates site exploration." There was no doubt from the Ames experiences that the presence of panospheric video improved the efficiency of telescience.

Additionally, the Ames team noted that the panospheric imagery was not sufficient in resolution to allow for complete evaluation and interpretation of the environment. The monochrome camera pair was exceedingly useful in examining obstacles, and in selecting where to aim the high-resolution camera pair. The human-replicate pan/tilt pair was credited with substantially improved interpretations of geologic features, including the discovery of a Jurassic era rock fossil bed. This discovery was a landmark, as it was the first time geologists used a surrogate robot to make such a discovery, and indicates that telexploration can in fact provide detailed understanding of a remote environment.

Perhaps the most concise evaluation of the Ames telescience experiments and the Atacama Desert Trek was spoken by Dave Lavery of NASA Headquarters: The combination of panospheric video and high-resolution imagery has been clearly demonstrated to enhance the efficiency and capabilities of remote science and robotic exploration [Wettergreen97].

Evaluation of Thesis Assertion

The intent of this research was to investigate the assertion that video conveying the entire surroundings of a teleoperated mobile robot enhances situational awareness needed to react to untraversable terrain while exploring.

The observations of both novice and experienced telexplorers showed that immersive video was beneficial in achieving progress and maintaining operator comfort during remote exploration. These findings were confirmed by the NASA Ames telescience experiments. The success of RoverTV and the general preference of the "foveal-immersive" display mode indicated that it is not necessary to immerse the viewer in imagery in order to provide these advantages. The principal complaint of the panospheric video innovated for telexploration during the Atacama Desert Trek was with respect to its resolution being lower than viewers were accustomed.

This research has demonstrated that the thesis assertion holds, and that video which conveys the surroundings of a remote robot does benefit the telexplorer in achieving awareness, avoidance, and if need be, extrication from untraversable terrain.

Chapter Summary

- Nearly 100 operators were observed driving the Nomad robot, in six display modes ranging in field of view and panning ability.
- Display was performed on an immersive theatre, high-resolution monitor, and television.
- Operators preferred to have access to peripheral information, but actually desired to have all imagery presented in the foveal region of their vision in a familiar orientation.
- Experienced operators appreciated the improvement of panospheric video over traditional teleoperative visualization, and novices complained on the resolution and color as compared to human vision and video games.
- Immersive video does enhance the operators ability to both avoid and extricate from untraversable terrain.

Chapter 5

Retrospections

The accomplishments of the research and arenas for future work are profiled, and **Retrospections** on the contributions lead to insights from the author.

This research accomplishes the innovation of panospheric video for remote operation of mobile robots. A panospheric sensor is developed and demonstrated on an actual robot performing telexploration. Immersive imaging as a means of achieving situational awareness is evaluated in varied display configurations.

Future work exists in upgrading the sensor, advancing the single-image-multiple-viewer capability, porting traditional machine vision algorithms to the immersive domain, and improving immersive displays. The sensor could benefit from continued advances in imaging technology and changes suggested during deployment in the desert. Enabling multiple viewers, even at distant locations from each other, to interact with the video as a team would further increase the utility of panospheric imaging for telexploration. With feature detection, algorithms such as range from motion flow or image sequence, obstacle avoidance, integrated positioning can be adapted from the machine vision library. Upgrading domed displays to directly accept panospheric imaging formats or developing immersive personal displays is essential to the continued growth of immersive imaging in robotics.

Along with adding a previously absent technology to the arsenal of the robotics designer, the contributions of the research stem from the ability of the robot and/or its operator to utilize panospheric technology to simultaneously view in all directions. Telexplorers do not need to plan a motion or pan/tilt sequence in order to achieve the desired views, and while passing an object the front, side, and back of the object are automatically imaged. Every operator can customize the display of imagery to their own liking, rather than having to adapt to the imagery of the robot.

Utilization of panospheric video for robotic telexploration is an appropriate solution, even though current technologies fall short of achieving the incarnation of human vision. In its initial deployment, panospheric technology was shown to have significant value over traditional methods of image acquisition. Solutions to any problem previously discarded due to complexity should be reexamined in light of technological advances. Even when immersive imagery is made equivalent to human vision, acquiring the imagery is only one part of replicating human visual-based perception.

Accomplishments

The presented research accomplished panospheric sensor innovation and demonstration along with evaluation of varied display formats and fields of view for robotic telexploration. The availability of video presenting the entire surroundings of the mobile robot enhanced the operator's confidence and allowed improved reactions to untraversable terrain.

Panospheric Video

A panospheric sensor was designed and implemented. Starting with initial design goals and basic theory, a sensor was envisioned to capture the entire surroundings at a remote location and verified in simulation. The implemented sensor was placed on board a mobile robot and utilized as the primary means of image acquisition from a remote environment.

Over one million images were successfully acquired and transmitted without any adjustments to the sensor. Even without significant attention paid to implementing a hands-free sensor, during the 45 day desert trek, the only adjustment needed was wiping champagne from the lens when the robot was sprayed upon completing the 200km milestone. Toward the end of the desert trek, it was noticeable that the mirror had shifted slightly away from alignment with the camera axis, but the resulting image did not adversely affect telexploration.

Over 100 novice drivers used panospheric video to traverse over 60km while exploring. Even with a multiple second time delay, telexplorers with no prior experience were able to participate successfully in controlling the robot at the remote site. Additionally, nearly 5000 telexplorers were involved in viewing the live video and determining which direction to view, what objects to focus upon, and discussing features of the terrain that were visible in the immersive imagery.

Evaluation of Display Modes

Telexploration was performed using video displayed on an immersive theatre, projection-screen, personal computer, and broadcast cable television. The video was displayed within seconds of acquisition in multiple formats at multiple locations. At each location, viewers were able to control how they viewed the imagery even when not in control of the robot.

On the immersive theatre, the video was displayed to the telexplorers in full immersive, simulated pan/tilt, fixed-forward, raw panospheric, and virtual-reality modes. While presented with peripheral imagery the telexplorers avoided trouble more easily, and were successful in reacting to and planning routes away from and around untraversable terrain. The view of choice provided peripheral information from the remote location in the foveal region of the telexplorer's vision (the virtual-reality style viewing mode).

While novices and experts disagreed about the quality of the telepresence, they were equally competent in operating the robot.

Future Work and Extensions

With the completion of the proposed research, it is clear that panospheric video and immersive display add value to teleoperation of robots. The logical extensions to the research are upgrading the sensor, increasing capabilities for multiple users, enabling autonomous robotics algorithms, and enhancing the display.

Sensor Upgrade

Better packaging and a smaller size is needed to reduce weight and enable more positioning options. The components required for the sensor: optic, camera, cables, computer, frame-grabber, and power supply filled a medium suitcase. To have utility to most robotic applications, this package needs to be reduced to the size of a shoe-box, and eventually to the size of a soda can. Beyond shrinking the size of the components, the packaging needs to be redesigned to provide strength and protection at a lighter weight.

The most glaring problem of the sensor was exactly that: the glare. The optic needs to be redesigned to reduce the amount of glare and eliminate the internal reflections caused by the various surfaces. Additionally, an optic should be implemented with the full sphere of acquisition rather than extended panoramic, and the distribution of pixels should be more even so that the resolution is nearly constant at any vector.

To achieve more satisfying video, a higher resolution camera is required, with higher frame-rate, and true color. To duplicate the 640x480 image from a traditional field of view camera, a panospheric camera should be about 3000x3000 pixels. While such mega-pixel cameras are available, they operate at tenths of Hertz and are grayscale instead of color.

Teleoperative Capability Upgrade

Because panospheric video captures imagery in all directions, multiple viewers can be accessing different parts of the image for different purposes. By adding a layer of networking and data sharing atop the imagery, telexplorers at distant locations can work together as a team, pooling their expertise. Information noted by one viewer could be attached to the imagery and available for other viewers. One viewer could also inform another viewer of where they were looking, or all users could switch between each other's current views. As the robot becomes more capable, planning and control strategies would be needed to handle commands from multiple telexplorers who are "independently" using the robot as a shared resource.

In bandwidth limited applications, compression enhancements would certainly help. However, it is also possible to allow the viewers to identify the regions they wish to view, and only send those selected windows at video rate, transmitting the bulk of the surroundings several times a minute. A viewer can select multiple windows, essentially turning a single panospheric sensor into an infinite number of pan/tilt cameras.

Machine Vision Algorithm Ports

By porting a simple machine vision technique such as feature detection and recognition to the panospheric image space, many additional capabilities for the panospheric sensor become possible. Tracking objects as they move about the robot, or as the robot moves about them, allows for visual servoing as the basis for commanding the robot, collision warning/avoidance, and map construction.

Using image sequences along with other sensors such as odometry and inclinometers allows for position estimation to be performed from landmarks and improves dead reckoning capabilities by allowing visual evaluation of the extent of slippage. Using motion flow techniques, or using position tags on previous images as a stereo baseline, range information can be reconstructed to enable scene reconstruction. When the robot passes by an object, information continues to be available unlike current stereo systems that can only gain value from objects that are in the overlapping region of coverage.

Multiple panospheric sensors could be arranged to perform stereo processing, with vertical baseline and equilateral triangle placements the most likely for success. The obvious problem is that the sensors appear in each other's views, but if placed far enough apart the occlusion is minimal. Vertical baseline allows for unobstructed panoramic stereo, with the exception of cables going to the top camera.

Display Enhancements

The simplest method for displaying panospheric imagery is to essentially back-project the acquired imagery through an identical optic onto a surrounding spherical screen. This technique has not been demonstrated, but in theory should work. It does require the projector optic to be located at the center of the sphere, and limits the area in which viewing can occur. A larger dome allows more people, but requires a stronger projector.

Partitioning the imagery to multiple projectors is the current standard for domed or large screen displays. The calibration required between the projectors to provide seamless display and identical color balancing makes the systems too expensive for most telexploration needs. However, the projectors can be placed nearly anywhere and the image simply reformatted to produce the correct projection. This means that there is plenty of room for viewers.

The ideal solution for immersive image display is a personal goggle that has a display surface duplicating the viewing regions of the eyes. Synching the display with head tracking systems would allow for a fully immersive experience, but until this technology increases motion-sickness, headaches, and eye fatigue makes this the least viable solution currently.

Contributions

By adding panospheric imaging to the robot and presenting immersive and peripheral information to the telexplorer, a higher state of awareness of the remote environment is achieved. This was demonstrated for the human through this thesis, but holds true for the robot itself once machine vision techniques are ported to this domain. The capability to view any or all directions simultaneously provides an unprecedented level of immediate access to the entire surroundings in a compact form.

Rather than planning a sequence of imagery views, or worrying about missing data while having to look ahead for obstacles, immersive acquisition enables continual progress instead of today's "step and wait" teleoperation. Automatically capturing video of objects as the robot passes by means that front, side, and back are seen, providing much greater insight into the nature of the object.

With multiple ways to display video, the teleoperation environment can be customized to fit each operator's preferences rather than forcing the operator to adapt to the robot. As a result, a single operator interface can be constructed that would be equally usable for many robots. The operator's preferences for operating any number of robots can be personalized in the same manner as that of virtual desktops in workstations. The result is that the learning curve to adjust to the particulars of the robots imaging systems and methods for switching camera views can be significantly reduced.

Robots can be created with a lower cost in terms of improved reliability, and decreased mass and cabling complexity. The panospheric sensor is inherently superior to a pan/tilt camera in every way other than resolution. However, for most pan/tilt applications high resolution imagery is not needed, and so a panospheric camera will suffice. Cutting the pan/tilt mechanisms increases the reliability of the robot (eliminating moving parts), and reduces the number of actuators that require power and control. In the case where a panospheric camera replaces a camera array, the savings in complexity of cabling and mass and image acquisition/transfer can make a substantial difference.

Insights

The rapid advances in technology have enabled a previously rejected solution (hyperwide field of view imaging) to be resurrected with resounding success. The panospheric sensor is a simple and elegant solution for the image acquisition needs of robots. However, it is only viable due to advances in imaging technology, processing capabilities, and display media over the past few years. People often search for new solutions to problems, and would be just as well served to review previous ideas that were "ahead of their time".

Although the innovated sensor and immersive telepresence environment were successful, the resulting system is nowhere near the capability of human vision. Eventually the technology for image acquisition and display will allow full immersive telexploration that is functionally equivalent (and possibly superior) to actually standing in the remote location and looking around. Nonetheless, as far as robotics is concerned, achieving the identical quality is a minor part in reproducing the capabilities of human visual-based perception. The primary challenge still lies in creating the robot that can look with comprehension into the world around it.

The lack of a true robotics industry means that robotic components are currently high-priced custom fabrications or are cobbled together using commonly available parts. The home computer industry has come into existence and boomed in the past two decades. For robots to become more capable and begin to enter the household, the robotics industry will likewise need to become a true presence. At this time, there are clear needs in the robotics research community without any industry to provide the needed components, regardless of whether they are in power, processing, sensors, mobility, or software.

When this work was proposed, immersive imagery and the panospheric sensor were relatively unused or unknown in robotics, but, in part through this work and the high profile demonstration in an extreme setting on board Nomad, the concept has become a deserved fixture on the robotics landscape.

Glossary

Following are terms related to this thesis. They appear in the text, or are of general significance.

Apollo missions NASA's late 1960's and early 1970's manned missons to explore the moon, ranging from orbital missions to lunar landings.

Atacama Desert A barren desert in northern Chile, known as one of the most arid locations in the world. Some regions have received less than an inch of rain in the 20th century.

Atacama Desert Trek A field experiment conducted by Carnegie Mellon University's Field Robotics Center to demonstrate lunar relevant technologies. A teleoperated rover, Nomad, crossed over 200km of the Atacama Desert while being controlled from Pittsburgh and NASA Ames Research Center.

autonomous In the robotics world, the ability of a robot to perform tasks without human intervention or low-level command.

blooming An effect of CCD cameras in which the saturation of a particular pixel causes enhanced measurements to be seen in its neighboring pixels.

camera array Multiple cameras used to achieve a field of regard greater than the field of view of a single camera. For immersive imaging, the array is usually positioned with the cameras oriented to look outwards from a common centerpoint.

CCD camera A charge-coupled device camera which creates an image through the conversion of photons to charge potential.

Dante Robot developed by Carnegie Mellon University to descend into active volcanoes. The first Dante robot was deployed at Mt. Erebus in Antarctica and the second Dante robot was deployed at Mount Spurr in Alaska.

direct teleoperation Performance of teleoperation while actually on-board or attached to the robot. Driving automobiles and working with master/slave manipulator systems are common examples.

experienced operator A teleoperator who has previously performed with a similar robot under similar command/console/interface conditions.

field of view The region of the world captured by an imaging device. Most commonly expressed in degrees, but occasionally also in steradians or percent-field.

fish-eye optic An optic which produces a round image, usually covering a wide field of view. Typical fish-eye optics capture 170 to 180degrees (a full hemisphere).

fixed-forward mode A mode of displaying panospheric imagery during the observations in this research. Only a 60degree field of view region to the front of the robot was displayed, and there was no way to pan or tilt the view.

fixed-immersive mode A mode of displaying panospheric imagery during the observations in this research. A full 200degree field of view region centered at the front of the robot was displayed, and there was no way to pan or tilt the view.

foveal immersive mode A mode of displaying panospheric imagery during the observations in this research. A full 200degree field of view region was displayed using only 60degrees of the full screen, with pan/tilt ability enabled.

foveal views The portion of biological vision systems with high resolution. Humans have one central foveal region, some birds have two foveal regions in each eye (side and forward views).

immersive When pertaining to imagery, implying imagery which captures the full surroundings of the camera rather than merely one direction. When pertaining to displays, implying filling the entire field of view of the audience.

inside-out teleoperation Operating a robot in a remote location while having only data delivered from the robot to aid in operations.

Kilauea Volcanic lava-field in Hawaii used for planetary relevant robot experiments.

LunaQuest An edutainment based teleoperated-robot mission to the lunar surface involving a long distance, long duration tour of historic and geologic sites.

Lunokhod One of two Soviet teleoperated rovers which were deployed on the lunar surface in the early 1970's and were operated for a combined total of about 40km.

Marsokhod Soviet prototype martian exploration robot utilized for terrestrial telescience operations by NASA Ames Research Center.

Nomad Robot developed by Carnegie Mellon University's Field Robotics Center to serve as a platform for testing lunar relevant technologies during the Atacama Desert Trek.

novice operator A teleoperator of a robot who has no prior experience with either the specific robot being teleoperated, or with teleoperation of robots in general.

odometry Measurement and log of data from wheel rotation history on a mobile robot. Sometimes taken to mean any motion data used for dead-reckoning based positioning.

operator Someone who commands a robot.

outside-in teleoperation Teleoperation of a mobile robot from a distance while being able to actually see the robot from the operation console. The operator is able to use their own perceptions of the robot's environment to perform tasks.

overhead immersive mode A mode of displaying panospheric imagery during the observations in this research. A top-down view (hemispheric) of the robot and its surroundings was displayed using a 60degree region of the theatre screen. panning the imagery was enabled and resulted in spinning the imagery.

pan and tilt Panning refers to changing the direction of view in the azumuthal or horizontal direction. Tilting refers to changing the direction of view in the elevation or vertical direction.

pan-forward mode A mode of displaying panospheric imagery during the observations in this research. A 60degree field of view region of the view about the robot was displayed. and operators were allowed to pan and tilt the imagery.

pan-immersive mode A mode of displaying panospheric imagery during the observations in this research. A full 200degree field of view region about the robot was displayed, and operators were allowed to pan and tilt the imagery.

panospheric imaging A technology which allows a substantially spherical field of view to be captured, digitally processed and presented to an observer in the form of a fully immersive spherical perspective image, in both still and video formats.

peripheral views The portion of biological vision in which imagery is acquired, although not with the resolution of a foveal region. It is also common for extreme peripheral vision to be grayscale rather than color.

pixel The picture elements which make up a digital image. Each element usually has a bit-depth indicating the number of levels of intensity which can be discriminated. Pixels can also be color specific, acquiring only light from one frequency. Typical color images are broken down into two dimensional arrays of red, green, and blue pixels.

Ranger Any of the instrument packages hard-landed on the lunar surface by NASA during the buildup to the Apollo program.

RATLER A four wheeled rocker-bogey suspension robot developed at Sandia National Labs and augmented by Carnegie Mellon University's Robotics Institute to serve as a testbed for safeguarded teleoperation.

reflective optic A part of a lens system which is mirror-like in purpose, causing incident light to be reflected.

refractive optic A part of a lens system which is glass-like in purpose, causing incident light to be refracted.

resolution When pertaining to images, the number of pixels in the horizontal and vertical fields of view. When pertaining to general vision, the ability to discriminate features at a distance.

roverTV Broadcast of Atacama Desert Trek video on a Pittsburgh cable channel with call-in viewers to the station able to operate the robot and control the imagery using their touch-tone phones.

SCAMP Robot created by the University of Maryland's Space System Laboratory to serve as a teleoperated underwater camera platform for the purposes of recording and monitoring neutral-bouyancy experiments.

situational awareness Information and knowledge necessary to understand the immediate physical surroundings and its dynamics.

Sojourner Teleoperated rover deployed on the Martian surface by NASA in 1997 as part of the Mars Pathfinder mission. Sojourner was developed at the Jet Propulsion Laboratory and based on the six-wheel rocker bogey suspension system refined in the Rocky series of robots.

spherical convex mirror The shape of the mirror utilized in the panospheric sensor innovated during this research. Somewhat erroneously commonly called a hemispheric mirror, the mirror has a constant radius of curvature and is first surface mirrored on the convex side.

steradians A unit of measurement commonly known as a solid angle. There are 4Pi steradians in a complete sphere.

stereoscopic imaging The process of acquiring depth and range information through the use of two or more images taken of the same subject area from two different, but known positions.

Surveyor Any of the instrument packages soft-landed on the lunar suface by NASA as part of the build-up to the Apollo program.

telemetry The stream of data returned from a robot to the operator for the purposes of providing feedback about the robot's state and sensor readings.

telexploration Using remotely deployed instrument packages for the purposes of exploring an unknown environment. Commonly created in the form of a mobile robot.

teleoperation The process of controlling a robot which is usually at a remote location from the operator.

telepresence Providing an audience with the sensation of being located at a remote site.

telescience Using remotely deployed instrument packages for the purpose of analyzing an distant environment.

traversable Terrain which can be successfully and safely crossed by a mobile robot. Untraversable terrain is that which poses a danger or a barrier to the robot.

votive Determining a decision based on votes.

References

Following are complete references to material cited in the text.

[Alexander90] Harold L. Alexander. "Experiments in Teleoperator and Autonomous Control of Space Robotic Vehicles", SPIE Vol. 1388 Mobile Robots V (1990). pp560.

[Almaula86] Haren Almaula. Summary of Advanced Ground Vehicle Technology (AGVT) Teleoperator Activities. Tech Report - Fort Knox. AGVT Concept Evaluation Program.

[Axelrod88] Tim Axelrod et al. "Wide Field of View Camera", Energy and Technology Review. Lawrence Livermore National Laboratories. Dec88.

[Ayres41] Waldemar Ayres. US. Patent 2244235, "Cycloramic Optical System. 1941.

[Bogner95] Stephen L. Bogner. "An Introduction to Panospheric Imaging". Proceedings of IEEE Systems Man and Cybernetics, 1995.

[Brumitt97] Barry Brumitt. Personal conversations and interactions with the UGV program. Carnegie Mellon University's Robotics Institute. 1997

[Disney96] Walt Disney Animation Studios. Web site documentation regarding the movie ToyStory "http://www.toystory.com"

[Eveleth76] R.G. Eveleth. "Experimental Remote Control Vehicle (RCV)", Tech Report - Fort Hood: TCATA Headquarters FM-262.

[Fong95] Terrence Fong, et al. "Operator Interfaces and Network-Based Participation for Dante II", Preprint copy for SAE 25th Intl Conf on Environmental Systems. July 1995.

[Freethy82] Ron Freethy. "How Birds Work - A guide to bird biology". Blandford Press, Poole Dorset. 1982.

[Glumm92] Monica M. Glumm et al. "A Study of the Effects of Lens Focal Length on Remote Driver Performance", Tech Report - Army Research Laboratory ARL-TR-25. Nov92.

[Gordon66] D.A. Gordon. "Experimental isolation of the driver's visual input", Human Factors, Vol 8. pp 129.

[Hardy39] Wilbur Hardy. US Patent 2176554, "Periscope" 1939.

[Heiken91] Grant H. Heiken, et al. "Lunar Source Book: a user's guide to the moon". Cambridge University Press 1991.

[Hunt96] Theresa Hunt. Personal conversations and interactions with the SCAMP robot. University of Maryland's Space Systems Laboratory. 1996.

[Juday91] Richard D. Juday. US Patent 5067019. "Programmable Remapper for Image Processing". 1991.

[Katragadda95] Lalietsh Katragadda, et. al. "A Robotic Expedition for Mass Interaction on the Moon". Proceedings of the 12th SSI-Princeton Conference: Space Manufacturing X - Pathways to the High Frontier. Barbara Faughnan, ed. 1995.

[Katragadda96] Lalitesh Katragadda, et al. "Rover Configuration for Entertainment-Based Lunar Excursion", Proceedings of the AIAA Robotics Technology Forum: Advanced Developments in Space Robotics, Madison, WI. 1996.

[Klarer93] Paul Klarer, et. al. "A Highly agile Mobility Chassis Design for a Robotic All-Terrain Lunar Exploration Rover (RATLER)." Technical Paper published by Sandia National Laboratory. 1993.

[Kodak97] Kodak Industrial Imaging Division. Web site documentation regarding digital cameras. "http://www.kodak.com".

[Krotkov94] Eric P. Kroktov, et al. "Lunar Rover Technology Demonstrations with Dante and Ratler." PRoceedings of the International Symposium on ARtificial Intelligence, Robotics and Automation for Space". Jet Propulsion Laboratory, Pasadena, CA. 1994

[Kuban94] Daniel P. Kuban et. al. US Patent 5359363. "Omnivew Motionless Camera Surveillance System" 1994.

[Kress88] Gary Kress et al. "Sensorimotor Requirements for Teleoperation", Tech Report - Corporate Technology Center, FMC Corporation Report R-6279. Santa Clara, CA. Dec88.

[Lucasfilm97] LucasFilm Ltd. Web site documentation on the Star Wars film trilogy. "http://www.starwars.com"

[Marshall61] A.J. Marshall. "Biology and Comparative Physiology of Birds", williams, New York and London. 1961.

[MacDonald59] J. D. MacDonald. "Instructions to Young Ornithologists - Bird Biology", London Museum Press Limited. 1959.

[Martin95] Martin Martin, "Robotic Navigation using Sonar and Evidence Grids". Technical Report, Carnegie Mellon University Robotics Institute, 1995.

[McCutchen91] David McCutchen, US Patent 5023725, "Method and Apparatus for Dodecahedral Imaging System"

[McGovern80] Douglas E. McGovern. "Experiences in Teleoperation of Land Vehicles", Tech Report - Sandia National Laboratories SAND87-1980c.

[McGovern87] Douglas E. McGovern. "Experiences in Teleoperation of Land Vehicles", Tech Report - Sandia National Laboratories SAND87-1908-UC-150.

[McGovern90] Douglas E. McGovern. "Experiences and Results in Teleoperation of Land Vehicles", Tech Report - Sandia National Laboratories SAND90-0299-UC-515.

[Miller86] Dwight P. Miller. "Evaluation of Vision Systems for Teleoperated Land Vehicles", IEEE Control Systems Magazine. Vol 8. #3 pp 37.

[NASA68] National Aeronautics and Space Administration, "Countdown to Impact. A history of the Ranger Program." Johnson Space Center. Houston, TX. 1968

[NASA75] National Aeronautics and Space Administration. "Apollo Program Summary Report", Johnson Space Center, Houston TX. 1968.

[Nalwa96] Vishvjit S. Nalwa. "A True Omni-Directional Viewer". Technical Paper published by AT&T Bell Laboratories. 1996.

[Nayar97] Shree Nayar. Personal conversations and web site documentation regarding Omnicam. "http://http://www.cs.columbia.edu/CAVE/omnicam/.index.html" and "http://www.cyclovision.com" 1997.

[Nikon97] Nikon Company, "Nikkor Lenses". Melville. New York. 1997.

[Pepper78] Ross L. Pepper et al. "Display System Variables Affecting Operator Performance in Undersea Vehicles and Work Systems", Tech Report - Naval Ocean Systems Center TR-269.

[Pepper86] Ross L. Pepper. "Human Factors in Remote Vehicle Control", Proc of the Human Factors Society - 30th Annual Meeting. pp 417. 1986.

[Southwell96] David Southwell, et al. "Panoramic Stereo", Proceedings of the International Conference on Pattern Recognition. Vienna, Austria. 1996

[Thomas98] Geb Thomas. "Reeal-time Panospheric Image Dewarping and Presentation for Remote Mobile Robot Control. Submitted to the IEEE Transactions on Robotics and Automation. 1998.

[Wettergreen97] David Wettegreen, et.al. "Operating Nomad during the Atacama Desert Trek", Technical publication of NASA Ames Research Center, Intelligent Mechanisms Group. 1997.

[Whittaker9X] William Whittaker, et. al. Technical Reports and Documents of the Lunar Rover Initiative. Carnegie Mellon University. Robotics Institute. 1993-8.

[Whittaker98] William Whittaker, et. al. "The Atacama Desert Trek: Outcomes". Proceedings of the International Conference on Robotics and Automation, Belgium 1998.