N94- 34031

CONTROLLING TELEROBOTS WITH VIDEO DATA AND COMPENSATING FOR TIME-DELAYED VIDEO USING OMNIVIEW™

Dan Kuban, Steve Zimmermann & Lee Martin
TeleRobotics International, Inc., 7325 Oak Ridge Highway, Knoxville, TN 37931

ABSTRACT

Remote viewing is critical for teleoperations, but the inherent limitations of standard video reduce the operator's effectiveness. These limitations have been compensated for in many ways, from using the operator's adaptability, to augmenting his capability with feedback from a variety of sensors and simulations. Omniview™ can overcome some of these limitations and improve the operator's efficiency without adding additional sensors or computational burden. It can minimize the potential collisions with facility equipment, provide peripheral vision, and display multiple images simultaneously from a single input device. The Omniview™ technology provides electronic pan, tilt, magnify, and rotational orientation within a hemispherical field-of-view without any moving parts. Image sizes, viewing directions, scale and offset etc., may be adjusted to fit operator needs.

This paper discusses the derivation of the image transformation, the design of the electronics, and two applications to telepresence that are under development. These are Video Emulated Tweening (VET), and Manipulator Guidance and Positioning (ManGAP). The VET effort uses Omniview™ to compensate for time-delayed video in teleoperation of remote vehicles. In ManGAP two Omniview™ systems are used to provide two sets of orientation vectors to points in the field-of-view (FOV). These vectors then provide absolute position information to both control the position of a telerobot, and to avoid collisions with the work sight equipment.

INTRODUCTION

Remote viewing is the most critical feedback in teleoperations. Close viewing is necessary for detailed manipulation tasks, while wide-angle viewing aids the positioning of the remote handling system and helps avoid collisions in the work space. The majority of these systems use either a fixed-mounted camera with a limited viewing field, or they utilize mechanical pan-and-tilt platforms and mechanized zoom lenses to orient the camera and magnify its image. These mechanisms can be large, unreliable, and may interfere or collide with the environment. Also, several cameras may be necessary to provide wide-angle viewing or complete coverage of the work space. Camera viewing systems that use prisms or mirrors to provide wide viewing angles have been developed in order to minimize the size and volume of the camera and minimize the amount of intrusion into the viewing environment, but this approach can result in blind spots. Also, these systems typically have no means of magnifying the image and or producing multiple images from a single camera.

The Omniview™ solution is based on the property that a fisheye lens allows a complete hemispherical field-of-view to be captured, but with significant barrel distortion present in the image periphery. A high speed image transformation processor has been developed that reconstitutes portions of the image to correct the lens distortion for display on an RS-170 standard format monitor. The Omniview™ imaging system has several advantages over standard camera systems. Multiple images may be simultaneously produced by the device allowing a single omnidirectional camera to provide numerous independent views from one location. The transformation is accomplished electronically, providing complete programmable control over viewing parameters.

IMAGE TRANSFORMATION

The postulates and equations for transforming the input image are based on the camera system utilizing a fisheye lens as the optical element. There are two basic properties and two basic postulates that describe the perfect fisheye lens system. The first property of a fisheye lens is that it encompasses a 2π steradian or hemispherical field-of-view and the image that it produces is a circle. The second property of the lens is that all objects in its field-of-view are in focus, i.e. the perfect fisheye lens has an infinite depth-of-field. In addition to these two main properties, the two important postulates of the fisheye lens system are stated as follows:

Postulate 1: Azimuth angle invariability - For object points that lie in a content plane that is perpendicular to the image plane and passes through the image plane origin, all such points are mapped as image points onto the line of intersection between the image plane and the content plane, i.e., along a radial line. The azimuth angle of the image points is therefore invariant to elevation and object distance changes within the content plane.

Postulate 2: Equidistant Projection Rule - The radial distance, r, from the image plane origin along the azimuth angle containing the projection of the object point is linearly proportional to the zenith angle ß, where ß is defined as the angle between a perpendicular line through the image plane origin and the line from the image plane origin to the object point.

Using these properties and postulates, the mathematical transformation for obtaining a corrected perspective image can be determined. These have been reported previously.[1] By knowing the desired zenith, azimuth, and object plane rotation angles and the magnification, the corrections to the input image can be calculated. This relationship provides a means to transform an image from an input image memory buffer to an output image memory buffer exactly. Also, the fisheye image system is completely symmetrical about the zenith; therefore, the vector assignments and resulting signs of various components can be chosen to reflect the desired orientation of the object plane with respect to the image plane. In addition, the transformation can be modified for various lens elements as necessary for other fields-of-view.

SYSTEM DESCRIPTION

The system consists of a wide angle lens, camera, Omniview™ transformer, display controller, and video monitor. The system is designed to be independent of the camera/lens and monitor and can be used with CCD or tube cameras, visible or infrared spectrums.

A block diagram of the prototype system is shown in Figure 1. The camera input image capture electronics uses a parallel RS-485 type interface to capture the output of the camera. The input and output image memory buffers consist of video RAM arrays with 8 bit resolution. The output display electronics provides a gray-scale 60 Hz interlaced display for an RS-170 standard display monitor. The 80C196 core provides the control interface functions for the prototype system as well as the calculation of the coefficients and parameters for the image transformation core. The trigonometric functions (sin,cos,tan) were implemented using a lookup table with resolution to within a degree. This was found to be sufficient since the direction-of-view parameters are input to the camera system as direct angles for pan, tilt, and rotation. There are two independent processor channels that calculate the corrected pixel positions corresponding to the mapped input coordinates for each direction-of-view. The image transformation processor is pipelined using both high speed arithmetic devices and FPGA elements in order to maximize overall performance.

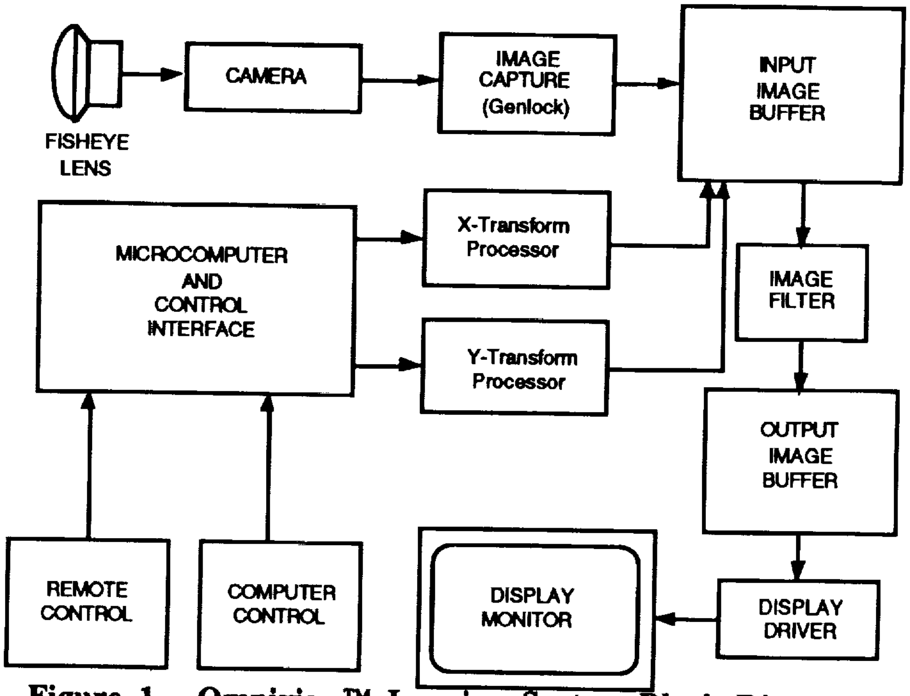

Figure 1 - Omniview™ Imaging System Block Diagram

APPLICATION OF OMNIVIEW™ TO TELEPRESENCE

The Omniview™ technology has many applications in remote viewing and telepresence. Two such development activities are currently underway at TRI. These are Video Emulated Tweening (VET), and Manipulator Guidance and Positioning (ManGAP). The VET effort uses Omniview™ to compensate for time-delayed video in teleoperation of remote vehicles. It relies on Omniview™'s capability to reorient the image without moving the camera to provide the operator with virtual video frames in between the real frames.

VET

A number of space related and teleoperated activities involve the transmission of slow-scan images (image updates slower than the standard 30 frames per second) due to transmission bandwidth or distance. When the slow-scan image is combined with direct operator interaction (for moving a vehicle, for manipulating an object, or for docking two vehicles) the operator often has to employ a "move and wait" strategy to overcome the delays associated with the video update. Of the methods used to counteract this problem, the most common approach presently under development involves predictive graphic simulation of the environment for projecting future actions.

The Omniview™ provides the ability to reorient the camera image without any motion of the camera or its video output, giving the operator the perception that the camera is moving. In practice, the perception of motion can be generated by modifying the pointing angle or magnification values in the transformation. Panning and tilting the image emulates turning and climbing, while magnifying and rotating emulates forward motion and tipping. For example, by matching the vehicle forward speed to the magnification, the operator can perceive vehicle motion by only manipulating the video image. This virtual motion has been demonstrated by using an enlarged aerial photo to simulate flight. The VET seeks to unite this perception of motion with the vehicle characteristics to provide an accurate and realistic emulation of continuous vehicle teleoperation with time delayed-video.

VET creates real-time intermediate video frames from live slow scan video based on vehicular motion commands. Utilizing slow scan video input, it captures the most recent image and adjusts the perspective in real time based on drive commands to the vehicle. The prototype system provides the operator with 22 frames/sec video yielding the perception of non-delayed communications through virtual video during the "delay interval".

This objective has been demonstrated on a vehicular viewing/operation system using Omniview™ with a slow scan video input (1 frame every four seconds) and vehicle control inputs to control the pan, and zoom of the image. The video camera is mounted on the remotely controlled vehicle. Figure 2 shows the image updates source verses time for real time video and for simulated video. At each live interval, a new video frame is captured and displayed, but some 100-140 intermediate frames are generated in between these real frames. For the ground vehicle demonstration development (a radio controlled car), two parameters were varied. The forward vehicular motion was simulated by zooming the image, and the turning of the vehicle was simulated by panning the image.

The match between the last simulated frame and the next live frame must be reasonable to insure that the operator does not receive a disturbing discontinuity in the displayed video. Live full frame video from the moving camera was recorded and compared to that produced with the slow-scan video and VET. Comparison of these two video results indicates that the degree of matching performed by the simulation relative to the actual image is sufficient to convince the operator that he has continuous motion. The effort surpasses a graphically generated approach by using the actual video image as the foundation for the tweening simulation of the remote vehicle, without the computational burden associated with graphic manipulation. The results are not only applicable to remote vehicular operation, but also to robotic teleoperation and spacecraft docking maneuvers.

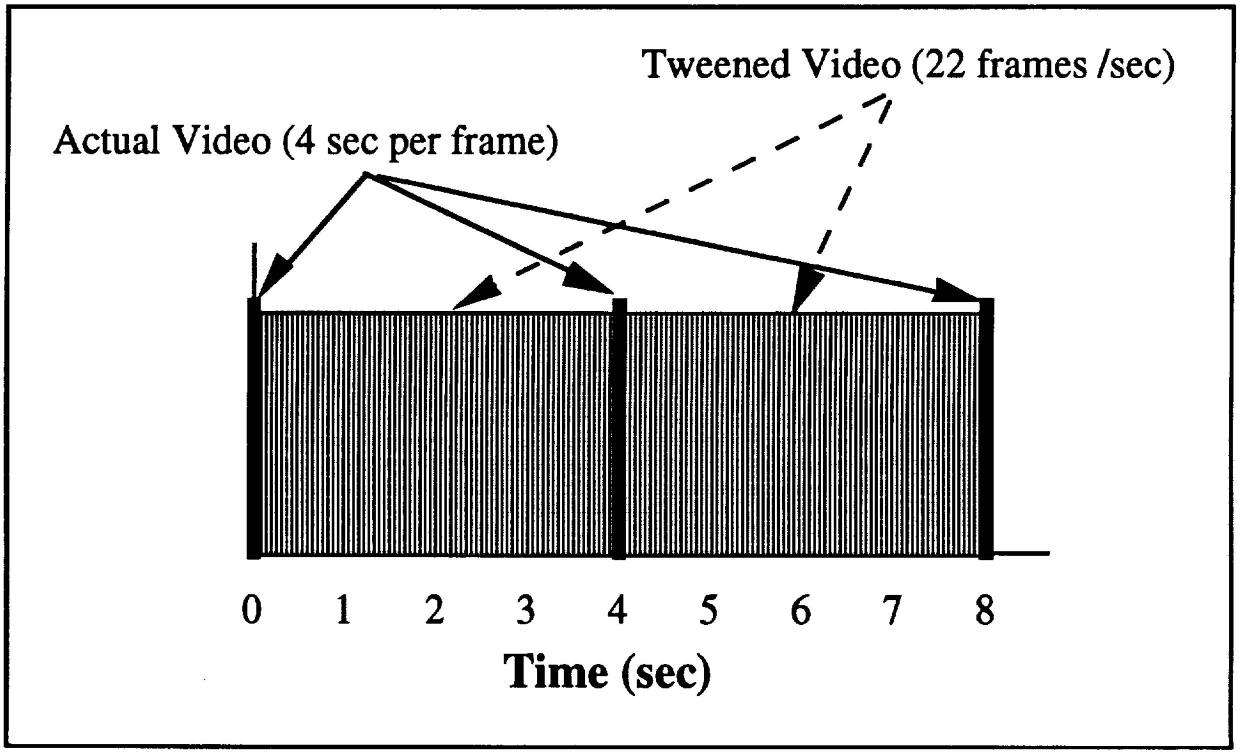

Figure 2 - Image update versus source for transmitted image and emulated image.

System block diagram is in Figure 3. The user interface and simulation subsystems obtain commands from the operator to control the vehicle, and then use those same commands to model the vehicular motions. Two information paths are initiated by the vehicle radio control transmitter. One path controls the vehicle via an RC link. The second RC link path actuates servos that allow the commands that are being sent to the vehicle to be monitored and read by the simulation subsystem. This second RC path receives the control signal, drives servos similar to the ones on the vehicle, converts the mechanical movement of the servo to an electric voltage via a linear potentiometer, and then samples this voltage using an analog to digital converter.

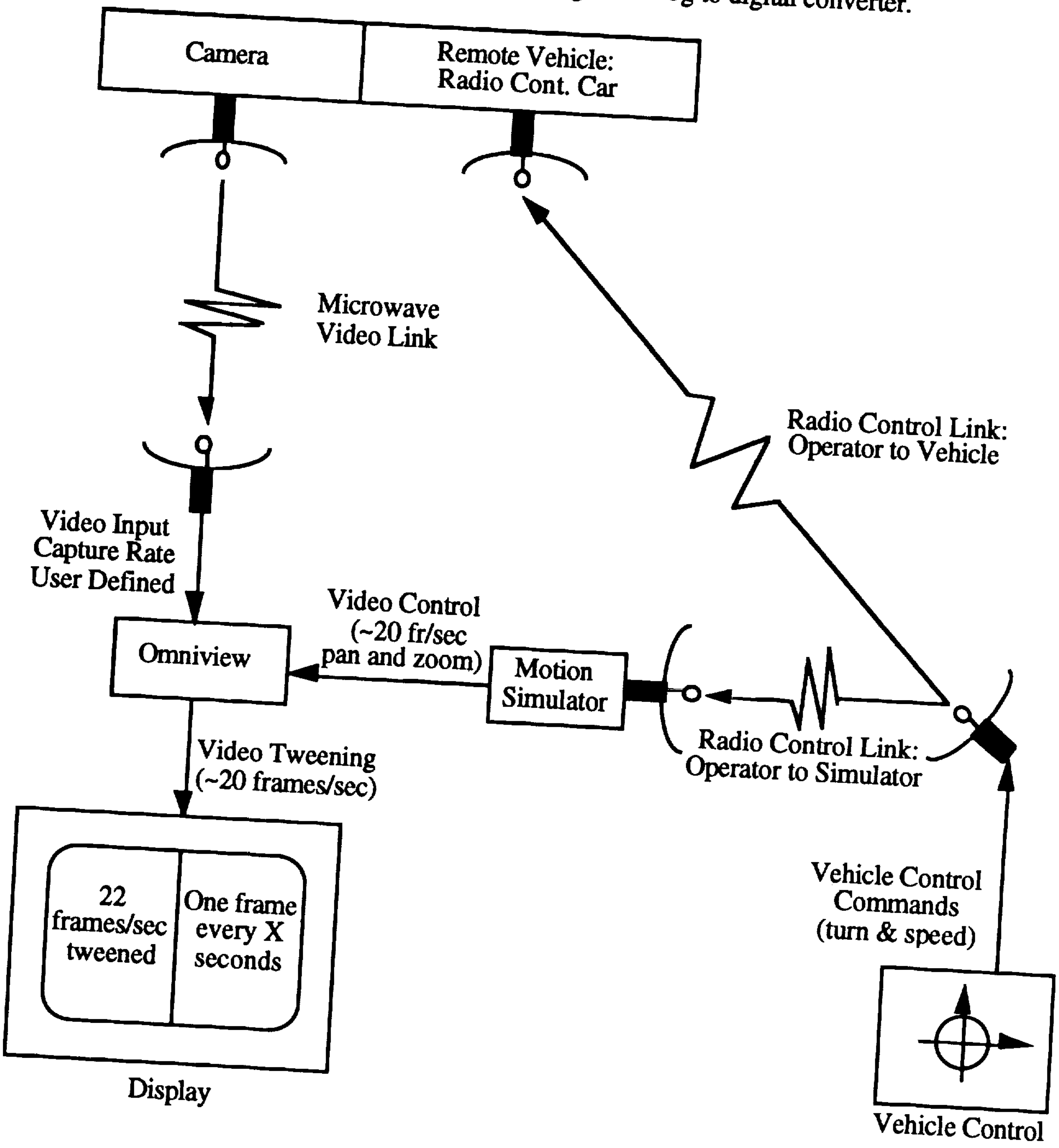

Figure 3 - Block Diagram of Phase 1 Implementation.

The user interface has been configured so that the key simulation parameters can be modified from a series of input switches. Using these switch inputs with observations of the vehicle and the images from it, an opportunity to empirically tune the video tweening model is provided during vehicle operation. For example, the effect of the zoom gain is to make forward motion appear to be occurring even though the input video is frozen due to slow scan time delays. As the zoom gain is increased, the vehicle appears to be moving at a higher rate of speed. At some point, the future simulated image and present actual image will converge. If the zoom gain is too low, the transition from future simulated to present actual images will appear to lurch forward for an instant. Conversely, if the zoom gain is too high, the simulated vehicle appears to move faster than the actual vehicle causing a reverse jump at the transition. If the magnification rate is matched then there is a smooth transition from the last virtual image to the next real frame, achieving the desired VET effect.

It works! In a very qualitative sense, Video Emulated Tweening (VET) achieves virtual reality. It gives the operator the perception of motion even though a still image is all that is available as input to the system. The transition between last virtual frame and first new image was not totally seamless in the prototype implementation, but the possibility for seamless performance exists if a reasonable knowledge of the relationship between the video source and the principle objects in the field are known, and if interlacing effects are eliminated through further development of Omniview™.

ManGAP

The second telepresence development activity that takes advantage of Omniview is the Manipulator Guidance and Protection (ManGAP) system. This system will be implemented and tested as part of the Integral Fast Reactor Program at the Argonne National Lab- West (ANL-W). This effort is driven by experience from operating remote facilities that has shown that transporting and positioning of remote handling equipment typically requires in excess of 50% of the total task completion time. The ManGAP applies video data from Omniview™ to minimize operator effort in the positioning of the teleoperator and transporter system.

The Omniview™ transformation is based on the orientation vector of the direction of interest relative to the camera axis. The three orientation angles of pan, tilt, and rotation to any point in the field-of-view are available from the Omniview™ processor. By selecting a point in the field-of-view on the monitor, the three orientation angles to this point relative to the camera axis are known. If a second Omniview™ is used and offset from the first, and the same point in the field-of-view is selected on the second monitor, then a second set of orientation angles is known. With a fixed offset between the two Omniview™ cameras, these two orientation vectors can be used to triangulate the X,Y,Z position of the selected point relative to the cameras. For controlling a teleoperator, the system requires a fixed location of the teleoperator relative to the Omniview™ pair, and the inverse kinematic transformations for the arm.

In the ManGAP system two Omniviews™ will provide plan view and front view coverage. The operator will use the front view to select the destination of the next motion, and utilize the plan view to designate the distance to the ending location. In this way, the operator will be able to fly the teleoperator end-effector or transporter to an end location by simply selecting the destination on two monitors. The ManGAP block diagram is shown in Figure 4.

A second realm of operation is also being developed - video based collision avoidance. In this mode the control system determines the direction for movement, redirects the manipulator along the line of movement and initiates a sequence of motion constraints to minimize the potential of collision between the manipulator and the working environment.

In this mode the operator determines a geometric area on the first monitor by drawing a graphical square, rectangle or circle. In the second monitor, he indicates a depth, thereby determining a volumetric boundary or envelope. With the Omniview™ orientation vector data, the location and size of this envelope are determined. The operator can define this envelope as a "keep-out" zone, or a "safe" zone. The ManGAP control system will then constrain the teleoperator to "stay out of" or "stay inside" this geometric envelope. This provides a level of collision avoidance and protection to equipment, but is not fully autonomous. It relies on the operator's intelligence to define the envelope. As such, it is a transition capability between teleoperation and total autonomy, combining human intelligence and machine control.

Overall, the main advantage of ManGAP is the ability to provide robotic control and collision avoidance without any additional sensors (and their associated cabling and control hardware). It simply uses the video data that is already present in any telepresence system.

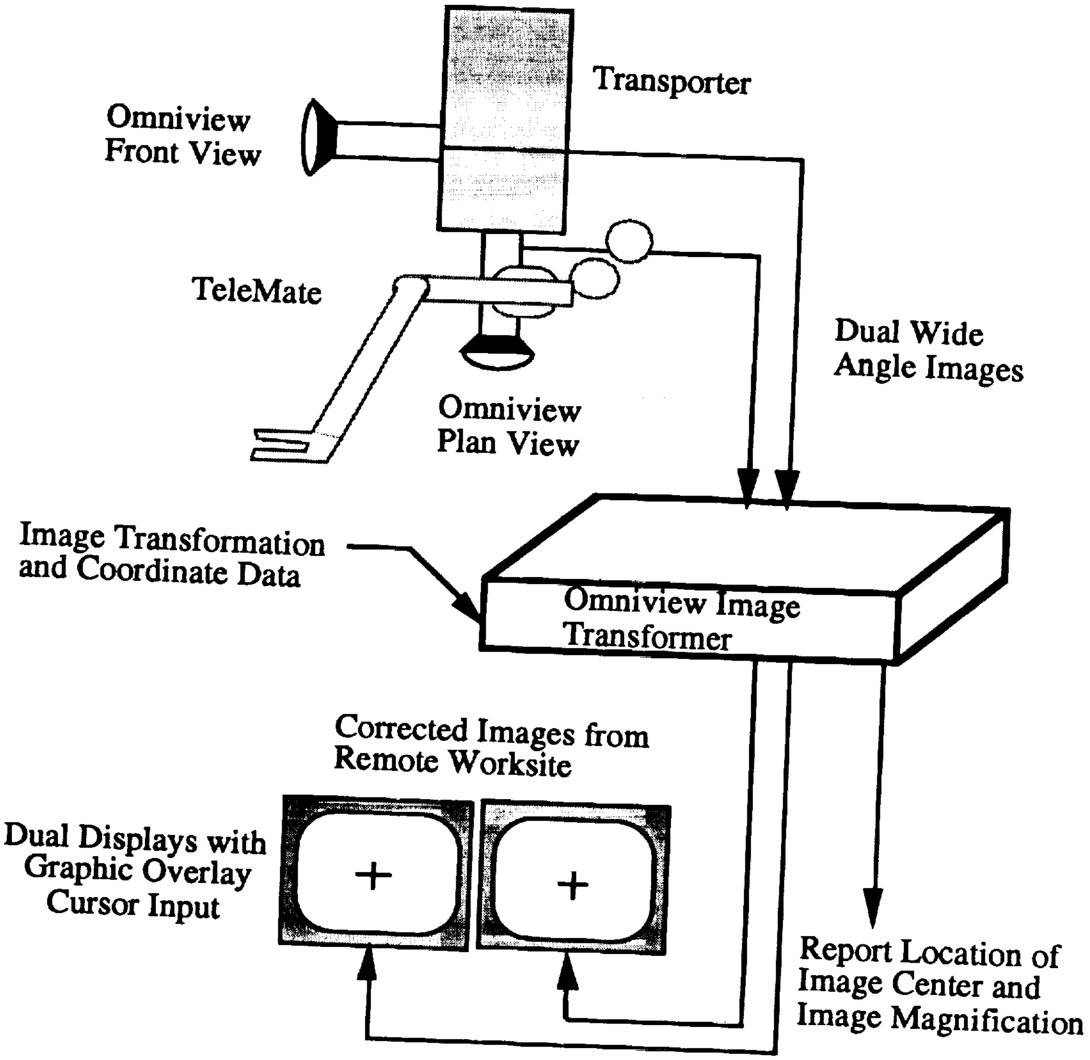

Figure 4 - Manipulator guidance and protection system hardware block diagram.

SUMMARY

Omniview™'s unique capabilities provide significant advantages in teleoperation and virtual environments. The feasibility of telerobot position control and collision avoidance using only video data promises to simplify telerobotic implementations by reducing sensors, cabling and computational requirements. It can also form the basis for an effective compensation of time-delayed video in teleoperations. The real-time demonstration of video manipulation yields convincing proof of virtual motion. This can improve the efficiency of teleoperations as well as provide alternatives to predictive graphical models.

REFERENCE

[1] Zimmermann, S.D. and Martin, H.L., "A Remote Video Pan/Tilt/Magnify/Rotate System with No Moving Parts", 1992 ANS/ENS International Meeting, Nov. 1992.

Session R6: MANIPULATORS AND END EFFECTORS

Session Chair: Mr. Charlie Price