

AN ELECTRONIC PAN/TILT/ZOOM CAMERA SYSTEM<sup>†</sup>

Steve Zimmermann  
TeleRobotics International, Inc.  
7325 Oak Ridge Highway  
Knoxville, TN 37921

H. Lee Martin  
TeleRobotics International, Inc.  
7325 Oak Ridge Highway  
Knoxville, TN 37921

## ABSTRACT

A camera system for omnidirectional image viewing applications that provides pan, tilt, zoom, and rotational orientation within a hemispherical field-of-view utilizing no moving parts has been developed. The imaging device is based on the effect that the image from a fisheye lens, which produces a circular image of an entire hemispherical field-of-view, can be mathematically corrected using high speed electronic circuitry. More specifically, an incoming fisheye image from any image acquisition source is captured in memory of the device, a transformation is performed for the viewing region-of-interest and viewing direction, and a corrected image is output as a video image signal for viewing, recording, or analysis. As a result, this device can accomplish the functions of pan, tilt, rotation, and zoom throughout a hemispherical field-of-view without the need for any mechanical mechanisms. A programmable transformation processor provides flexible control over viewing situations. Multiple images, each with different image magnifications and pan-tilt-rotate parameters, can be obtained from a single camera. The image transformation device can provide corrected images at frame rates compatible with RS-170 standard video equipment. The device can be used for many applications where a conventional mechanical pan-and-tilt orientation mechanism might be considered including inspection, monitoring, surveillance, and target acquisition. Omniview is ideal for multiple target acquisition and image stabilization in military applications due to its multiple image handling and fast response capabilities.

## INTRODUCTION

Camera viewing systems are abundant in surveillance, inspection, security, and remote sensing applications. Remote viewing is critical for robotic manipulation tasks, and is often performed by cameras attached to mechanisms that provide the pan, tilt, zoom, and focus capabilities. Close viewing is necessary for detailed manipulation tasks while wide-angle viewing aids positioning of the robotic system to avoid collisions with 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 cause interference and collision with the environment. In the applications where orientation of the camera and magnification of its image are required, the mechanical solution is large and can subtend a significant volume making the viewing system difficult to conceal or use in close quarters. Also, several cameras may be necessary to provide wide-angle viewing or complete coverage of the work space. Camera viewing systems that use internal optics 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. These systems rely on the movement of either a mirror or prism to change the tilt-angle of orientation and provide mechanical rotation of the entire camera to change the pitch angle of orientation. Using this approach, the size of the camera orientation system can be minimized, but "blind spots" in the center of the view result. Also, these systems typically have no means of magnifying the image and or producing multiple images from a single camera.

In order to minimize the size of the camera and orientation mechanism, a camera system was developed for remote viewing applications that utilizes fisheye optics and electronics processing to provide pan, tilt, zoom,

<sup>†</sup>Work sponsored by NASA Langley Research Center under contract NAS1-18855.

and rotational capabilities within a hemispherical field-of-view with no moving parts. The Omniview camera approach is based on the property of a fisheye lens which 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 with the correct perspective for display on an RS-170 standard format monitor. The Omniview camera system has several advantages over other camera systems. The implementation is such that multiple images may be simultaneously produced by the device allowing a single omnidirectional camera to provide numerous views from one location. The transformation is accomplished electronically, providing complete programmable control over viewing parameters. Image sizes, viewing directions, scale and offset etc. may be adjusted to fit operator needs. Since the fisheye image is symmetrical about the image center the camera need not be oriented vertically and can be rotated to match operator.

Potential applications of the Omniview system include remote viewing in constrained environments, inspection, object tracking, and surveillance applications. Since the Omniview camera system can produce multiple images from a single fixed camera, it can replace several camera systems in a remote viewing system.

## DEVELOPMENT OF THE MAPPING ALGORITHM

The postulates and equations that follow 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  $2\pi$  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  $\beta$ , where  $\beta$  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. Thus the relationship:

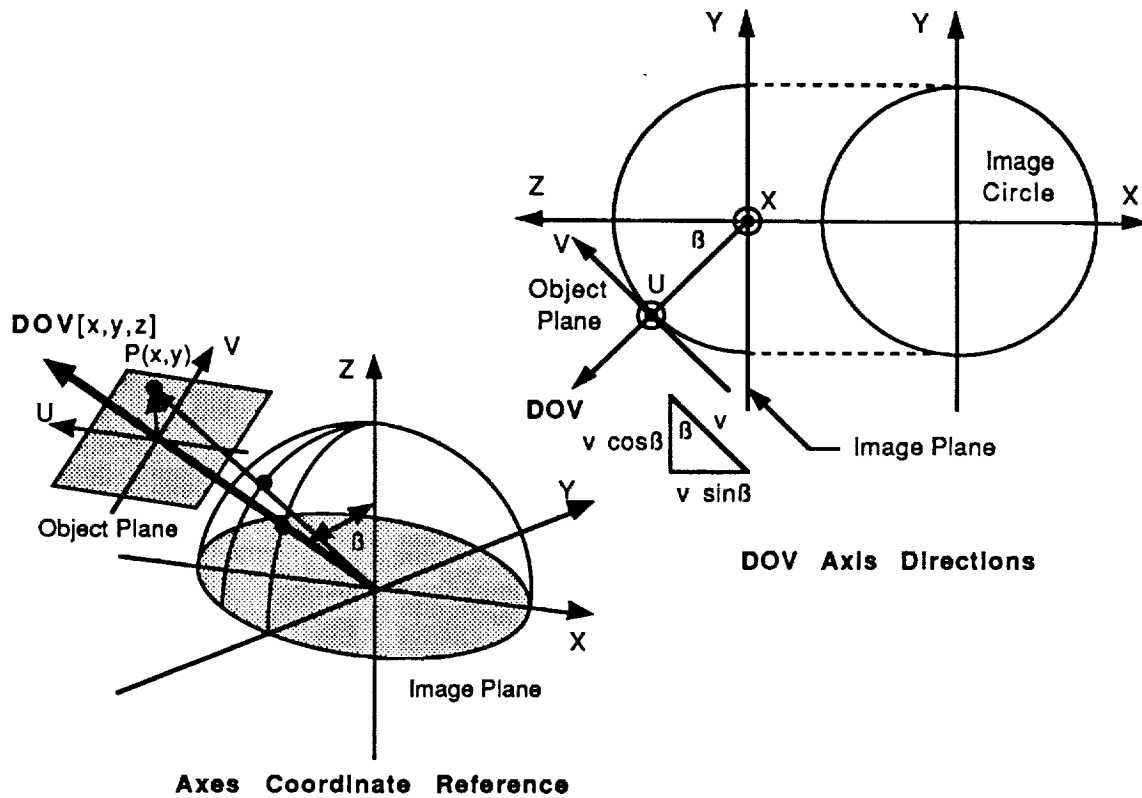
$$r = k\beta \quad (1)$$

Using these properties and postulates as the foundation of the fisheye lens system, the mathematical transformation for obtaining a perspective corrected image can easily be determined. The picture in Figure 1 shows the coordinate reference frames for the object plane and the image plane. The coordinates  $u,v$  describe object points within the object plane. The coordinates  $x,y,z$  describe points within the image coordinate frame-of-reference.

The object plane shown in Figure 1 is a typical region-of-interest that we desire to determine the mapping relationship onto the image plane to properly correct the perspective of the object. The direction-of-view vector,  $DOV[x,y,z]$ , determines the zenith and azimuth angles for mapping the object plane,  $UV$ , onto the image plane,  $XY$ . The object plane is defined to be perpendicular to the vector  $DOV[x,y,z]$ .

The location of the origin of the object plane in terms of the image plane coordinates is given by:

$$\begin{aligned} x &= D \sin\beta \cos\theta \\ y &= D \sin\beta \sin\theta \\ z &= D \cos\beta \end{aligned} \quad (2)$$



**Figure 1 - Coordinate Reference Frame Representation**

where  $D$  is the scalar length from the image plane origin to the object plane origin,  $\beta$  is the zenith angle, and  $\vartheta$  is the azimuth angle in image plane spherical coordinates. The origin of the object plane is represented as a vector using the components given in equation 1 as:

$$\mathbf{DOV}[x,y,z] = [D \sin \beta \cos \vartheta, D \sin \beta \sin \vartheta, D \cos \beta] \quad (3)$$

$\mathbf{DOV}[x,y,z]$  is perpendicular to the object plane and its scalar magnitude  $D$  provides the distance to the object plane. By aligning the  $YZ$  plane with the direction of action of  $\mathbf{DOV}[x,y,z]$ , the azimuth angle  $\vartheta$  becomes either 90 or 270 degrees and therefore the  $x$  component becomes 0 resulting in the  $\mathbf{DOV}[x,y,z]$  coordinates:

$$\mathbf{DOV}[x,y,z] = [0, -D \sin \beta, D \cos \beta] \quad (4)$$

Referring now to Figure 1, the object point relative to the  $UV$  plane origin in coordinates relative to the origin of the image plane is given by the following:

$$\begin{aligned} x &= u \\ y &= v \cos \beta \\ z &= v \sin \beta \end{aligned} \quad (5)$$

therefore, the coordinates of a point  $P(u,v)$  that lies in the object plane can be represented as a vector  $P[x,y,z]$  in image plane coordinates:

$$P[x,y,z] = [u, v\cos\beta, v\sin\beta] \quad (6)$$

where  $P[x,y,z]$  describes the position of the object point in image coordinates relative to the origin of the UV plane. The object vector that describes the object point in image coordinates is then given by:

$$O[x, y, z] = DOV[x, y, z] + P[x, y, z] \quad (7)$$

$$O[x, y, z] = [u, v\cos\beta - D\sin\beta, v\sin\beta + D\cos\beta] \quad (8)$$

Projection onto a hemisphere of radius  $R$  attached to the image plane is determined by scaling the object vector  $O[x,y,z]$  to produce a surface vector  $S[x,y,z]$ :

$$S[x,y,z] = \frac{RO[x,y,z]}{|O[x,y,z]|} \quad (9)$$

By substituting for the components of  $O[x,y,z]$  the vector  $S[x,y,z]$  describing the image point mapping onto the hemisphere becomes:

$$S[x,y,z] = \frac{RO[u, (v\cos\beta - D\sin\beta), (v\sin\beta + D\cos\beta)]}{\sqrt{u^2 + (v\cos\beta - D\sin\beta)^2 + (v\sin\beta + D\cos\beta)^2}} \quad (10)$$

The denominator in the last equation represents the length or absolute value of the vector  $O[x,y,z]$  and can be simplified through algebraic and trigonometric manipulation to give:

$$S[x,y,z] = \frac{RO[u, (v\cos\beta - D\sin\beta), (v\sin\beta + D\cos\beta)]}{\sqrt{u^2 + v^2 + D^2}} \quad (11)$$

From equation 11, the mapping onto the two-dimensional image plane can be obtained for both  $x$  and  $y$  as:

$$x = \frac{Ru}{\sqrt{u^2 + v^2 + D^2}} \quad (12)$$

$$y = \frac{R(v\cos\beta - D\sin\beta)}{\sqrt{u^2 + v^2 + D^2}} \quad (13)$$

Additionally, the image plane center to object plane distance  $D$  can be represented in terms of the fisheye image circle radius  $R$  by the relation:

$$D = mR \quad (14)$$

where  $m$  represents the scale factor in radial units  $R$  from the image plane origin to the object plane origin. Substituting equation 14 into equations 12 and 13 provides a means for obtaining an effective scaling operation or magnification which can be used to provide an equivalent zoom operation.

$$x = \frac{Ru}{\sqrt{u^2 + v^2 + m^2 R^2}} \quad (15)$$

$$y = \frac{R(v\cos\beta - mR\sin\beta)}{\sqrt{u^2 + v^2 + m^2 R^2}} \quad (16)$$

Using the equations for two-dimensional rotation-of-axes for both the UV object plane and the XY image plane the last two equations can be further manipulated to provide a more general set of equations that provides for rotation within the image plane and rotation within the object plane.

$$x = \frac{R[uA - vB + mR\sin\beta\sin\partial]}{\sqrt{u^2 + v^2 + m^2 R^2}} \quad (17)$$

$$y = \frac{R[uC - vD - mR\sin\beta\cos\partial]}{\sqrt{u^2 + v^2 + m^2 R^2}} \quad (18)$$

where

$$\begin{aligned} A &= (\cos\partial\cos\delta - \sin\partial\sin\delta\cos\beta) \\ B &= (\sin\partial\cos\delta + \cos\partial\sin\delta\cos\beta) \\ C &= (\cos\partial\sin\delta + \sin\partial\cos\delta\cos\beta) \\ D &= (\sin\partial\sin\delta - \cos\partial\cos\delta\cos\beta) \end{aligned} \quad (19)$$

and where

$$\begin{aligned} R &= \text{radius of the image circle} \\ \beta &= \text{zenith angle} \\ \partial &= \text{Azimuth angle in image plane} \\ \delta &= \text{Object plane rotation angle} \\ m &= \text{Magnification} \\ u, v &= \text{object plane coordinates} \\ x, y &= \text{image plane coordinates} \end{aligned} \quad (20)$$

The two equations expressed in 17 and 18 provide a direct mapping from the UV space to the XY image space and provide the fundamental mathematical foundation for the omnidirectional viewing system with no moving parts. By knowing the desired zenith, azimuth, and object plane rotation angles and the magnification, the locations of x and y in the input image can be determined. This approach 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, these postulates and mathematical equations can be modified for various lens elements as necessary for the desired field-of-view coverage in a given application.

## DESCRIPTION OF PROTOTYPE CAMERA SYSTEM

The Omniview camera system electronics was implemented using a single wire-wrapped prototype board. A photograph of the Omniview prototype system is shown in Figure 2. A photograph of the prototype electronics board and enclosure are shown in Figure 3. The prototype electronics board was mounted in a 12 by 14 inch enclosure with a +5V and  $\pm 12V$  power supply. A connector was provided for interfacing to the Videk digital camera, a BNC for the monitor display, a DB-15 connector for remote computer control, and a DB-15 connector for the remote hand-held controller. A Videk Megaplug model camera system was used for the development system. The Videk camera was chosen because it had the highest resolution CCD element for a commercially available camera during the initial phase of the project. The Megaplug camera provides a 1320 by 1024 resolution image at up to five frames per second. The Videk camera provided a good platform to demonstrate the prototype system. The camera uses a Nikon F-mount lens adapter and provides a standard 35-

mm back focal plane distance. The input lens element consisted of a Nikon 8-mm fisheye lens and a Nikon lens reduction element. The reduction element was necessary to reduce the 23-mm diameter circle produced by the fisheye lens to match the 2/3-inch format CCD imager size used in the Videk camera. Since the camera provides only a 10 MHz pixel scan rate output, the operating frequency of the wire-wrap implementation was lowered thereby reducing timing constraints.

A block diagram of the prototype system is shown in Figure 4. The camera input image capture electronics consists of a parallel RS-485 type interface to capture the digital output of the Videk camera. The input image memory buffer consisted of a 2048 by 1024 element video RAM array with 8 bit resolution. The input memory buffer was designed using 16 Texas Instruments TMS44C251 1-Megabit video RAMs organized as 512 by 512 element by 4 bit-planes. The output image memory buffer consists of a 1024 by 512 element array with 8 bit resolution and consists of 4 of the TMS44C251 video RAMs. The output display electronics provides a gray-scale 60 Hz interlaced display for an RS-170 standard display monitor.

The microcomputer and control interface for the prototype consists of a simple N80C196 microcontroller host processor core with 64K bytes of RAM and 64K bytes of EPROM. The N80C196 microcontroller provides a great deal of integration and timing control within a single 68-pin PLCC package. It contains an 8-channel 10-bit A/D converter system, several hardware and software timing units, a high-speed input-output peripheral system for event control, a high-speed serial unit and a high speed arithmetic ALU. The 80C196 microcontroller has been used successfully on several embedded Robotics, equipment control, communications, and multimedia platforms. Originally designed for the automotive environment, it provides an excellent core for embedded designs.

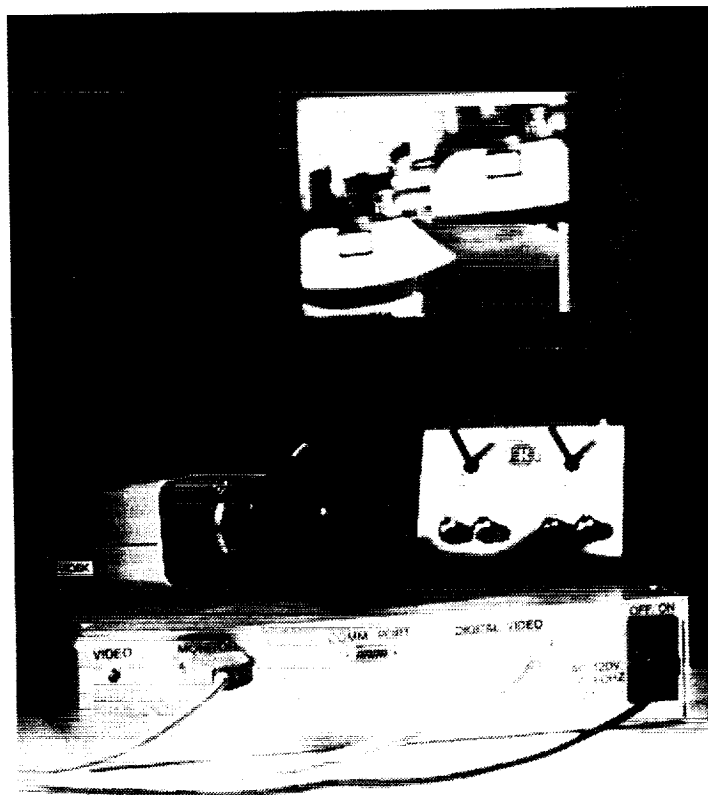


Figure 2 - Photograph of Omniview Camera System

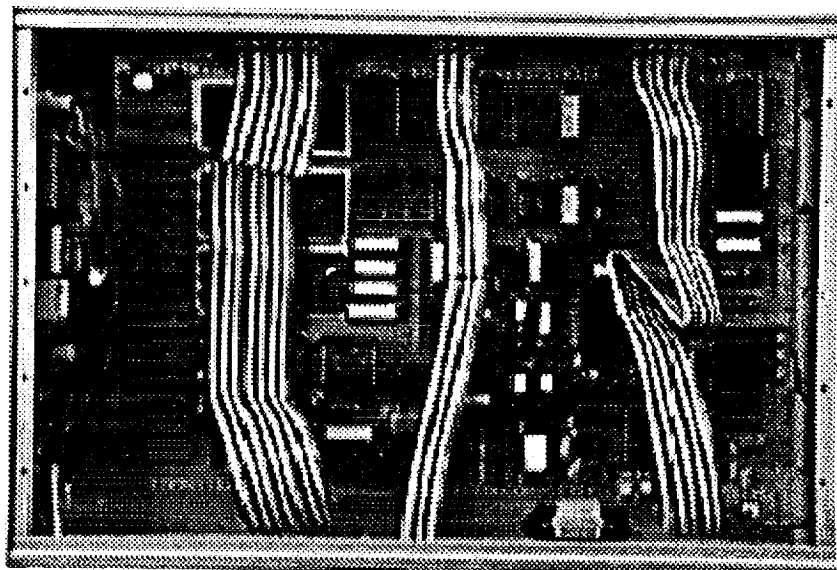


Figure 3 - Photograph of Prototype Electronics Enclosure

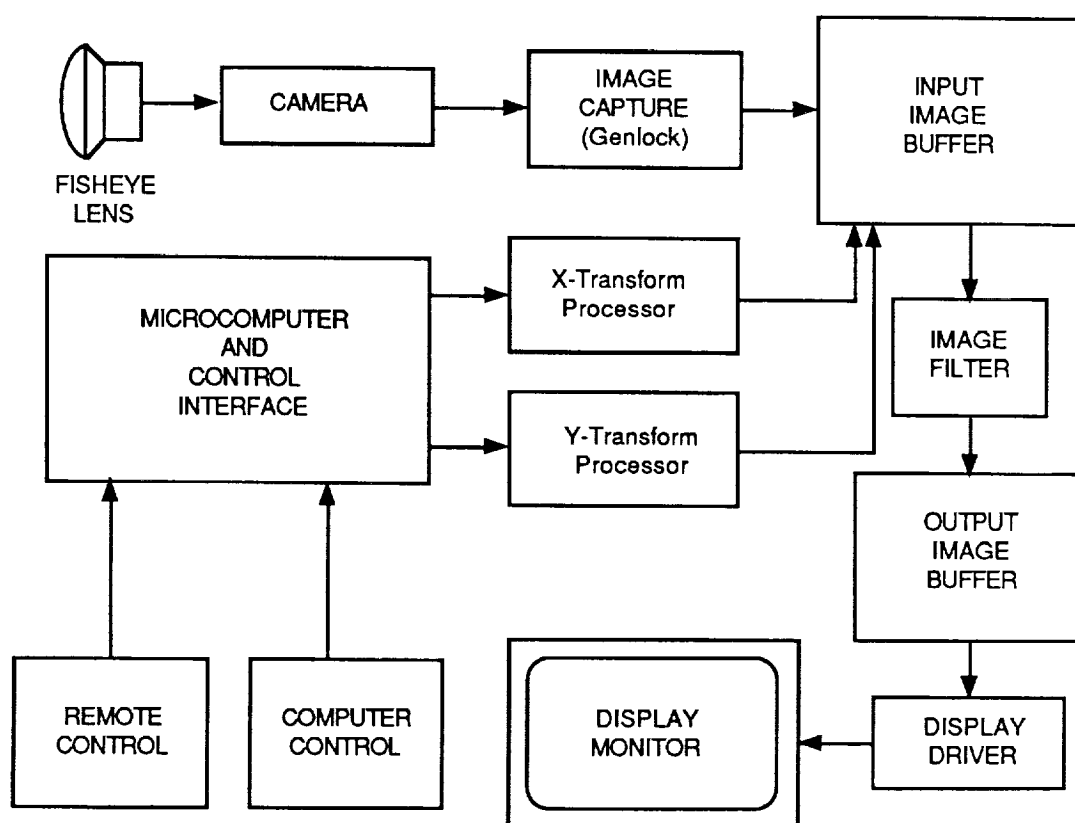


Figure 4 - Omniview Camera System Block Diagram

ORIGINAL PAGE  
BLACK AND WHITE PHOTOGRAPH

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. In order to support the coefficient calculations for the image processor core, a 96-bit and 48-bit arithmetic software package was implemented. 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. Also since the frame rate of the Videk camera was only 5 Hz, the calculational update rate of the host processor was sufficient to calculate a new set of parameters during each frame reception. A 48-bit precision square-root function was also implemented using Newton's approximation method.

The image transformation core and image filter consists of high-speed arithmetic devices that implement the basic transformation mapping as presented in equations 17 and 18. There are two independent processor channels that calculate the x and y pixel positions corresponding to the mapped u and v 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. A single programmable logic sequencer such as the Advanced Micro Devices AM29CPL154 handles all initialization and sequencing operations for the image processor system. The transform processor provides the capability to either nearest neighbor sample the input image space or to provide a 4-pixel bilinear filter on the image. A single 8-bit multiply-accumulate or MAC integrated circuit was used to implement the image filter. The image processor can be developed as a stand-alone core for use in other applications. The present design enhancements of the image transformation processor will allow it to be plugged into future designs.

A hand-held remote control interface was designed to provide an operator input device. The hand-held unit provides two x-y joystick interfaces for independent control of either one or two output images. Rotational potentiometers were provided for control of rotation offset and scale for each image. Using the hand-held unit in the dual image display mode, each image can independently be rotated, panned and tilted to the desired viewing angle. Two toggle switches on the front of the unit provide the capability to select between two different lens configurations and between one and two image display modes.

### Photographs of Prototype System Output

Several photographs of the output display of the prototype system in operation are shown in figures 5 through 8. The photograph in figure 5 shows a full 180-degree input image that is being captured by the fisheye lens mounted camera system. The output image is being displayed with no correction being applied. The photograph in figure 6 shows a view looking toward the right. The photograph in figure 7 shows a corrected image looking toward the left. Both corrected images are viewed at angles of up to 90 degrees from the line-of-sight of the camera. As can be seen from the images, the perspective has been corrected so that the barrel distortion evident in the input image has been removed. A good indication of the effect is the correction of the floor pattern and support pole in the fisheye image. The fisheye lens tends to produce circles out of straight lines for lines that are more distant from the center. Using this camera system, the viewing range contains the entire hemispherical field-of-view that is presented to the fisheye lens. One may look from the ceiling to the floor, a range covering 180 degrees, without moving the camera.

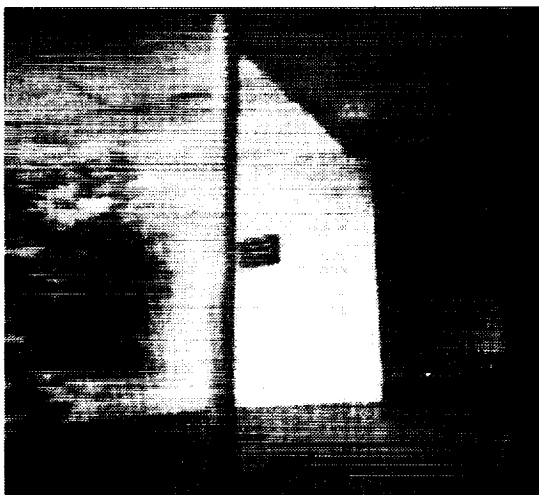
The photograph shown in figure 8 shows a simultaneous dual-image display using the same direction-of-view parameters for the images in figures 6 and 7. As shown, the Omniview camera system can display more than one simultaneous image from the same camera. Since the image sizes can be reduced to accommodate more than one image in the same pixel display area, the system is still capable of producing real-time updates with multiple images.



**Figure 5 - Fisheye Image**



**Figure 6 - Right View**



**Figure 7 - Left View**



**Figure 8 - Dual Left and Right View**

## SUMMARY

The Omniview camera system provides a unique solution for remote viewing requirements. A full hemispherical field-of-view capability provides a wider coverage than most mechanical pan-and-tilt mounted camera systems. The electronic implementation of the pan, tilt, rotation, and zoom functions provides a very flexible system to meet various viewing needs. Slewing from one direction-of-view to another direction-of-view can be accomplished frame-by-frame due to its programmable control of viewing parameters.

A complete prototype camera system has been implemented and demonstrated. The prototype camera system uses a high resolution CCD camera with a Nikon 8-mm fisheye lens and reducer arrangement for the input image capture device. The prototype system provides up to 5 frames per second image acquisition and up to 180-degree field-of-view. Various lens configurations can be used to achieve different field-of-view requirements.

A new design of the electronics is in progress at present. The new design will use an 80960SB 32-bit RISC processor for the host controller. The 80960SB provides a low-cost embedded controller core for high speed mathematical calculations. It contains a full IEEE 80-bit floating point core with support for trigonometric functions. The 80960SB provides enough bandwidth to support multiple image manipulation for a real-time 30 frame-per-second system. The new Omniview electronics system will also support up to 2048 by 2048 element input images for higher resolution viewing applications. The image core of the new design is essentially unchanged except for the integration of the control logic into pipelined FPGA architectures for faster memory access. The video RAM interface logic is being upgraded to support real-time image transformations.

Potential applications of the Omniview system include remote viewing in constrained environments, inspection, object tracking, and surveillance applications. Omniview is ideal for multiple target acquisition and image stabilization applications due to its multiple image handling and fast response capabilities. It can replace several camera systems in a remote viewing application.

The authors would like to express deep appreciation to Walter Hankins, Jack Pennington, and Al Meintel of the NASA Langley Research Center for their support in this research effort.

## REFERENCES

- [1] S. D. Zimmermann, Notes from Personal Journal.
- [2] S. D. Zimmermann, et. al., "Optimizing the Camera and Positioning System for TeleRobotics Worksite Viewing, Final Report for NASA Contract No.NAS1-18627.