

# Acquisition of stereo panoramas for display in VR environments

Richard A. Ainsworth\*, Ainsworth & Partners, Inc. (United States); Daniel J. Sandin, Univ. of California, San Diego (United States) and Univ. of Illinois at Chicago (United States); Jurgen P. Schulze, Andrew Prudhomme, Thomas A. DeFanti, Univ. of California, San Diego (United States); Madhusudhanan Srinivasan, King Abdullah University of Science and Technology (Saudi Arabia).

## ABSTRACT

Virtual reality systems are an excellent environment for stereo panorama displays. The acquisition and display methods described here combine high-resolution photography with surround vision and full stereo view in an immersive environment. This combination provides photographic stereo-panoramas for a variety of VR displays, including the StarCAVE<sup>[1]</sup>, NexCAVE<sup>[2]</sup>, and CORNEA<sup>[3]</sup>. The zero parallax point used in conventional panorama photography is also the center of horizontal and vertical rotation when creating photographs for stereo panoramas. The two photographically created images are displayed on a cylinder or a sphere. The radius from the viewer to the image is set at approximately 20 feet, or at the object of major interest.

A full stereo view is presented in all directions. The interocular distance, as seen from the viewer's perspective, displaces the two spherical images horizontally. This presents correct stereo separation in whatever direction the viewer is looking, even up and down. Objects at infinity will move with the viewer, contributing to an immersive experience. Stereo panoramas created with this acquisition and display technique can be applied without modification to a large array of VR devices having different screen arrangements and different VR libraries.

**Keywords:** stereo panorama, environmental imaging, VR, StarCAVE, NexCAVE, tiled displays, CORNEA, COVISE, KAUST, virtual reality.

## 1. INTRODUCTION

The dual-camera image capture system was developed to explore the combination of high-resolution photography and stereo panoramas with a variety of VR displays. Instead of a single camera, a pair of cameras, separated by the interocular distance, is used to capture two separate sets of images, each corresponding to the view of one eye. This camera pair rotates as a unit both horizontally and vertically, maintaining the center of rotation between the zero parallax points of the two lenses.

The focal length used determines the number of images required to cover a complete sphere and, hence, the resolution of the final result. In practice, the total resolution of the final image is typically set to at least double the resolution of the display system.

For the spherical imaging examples in this paper, a focal length of 40mm (35mm equivalent) was used. With this focal length, total image capture for a 360° x 180° stereo panorama typically requires 150 images and runs 1.8 gigapixels. After blending with the stitching software, the two final spherical images provide 800 megapixels of VR stereo information. For the rectangular display, the focal length was extended to 70mm (35mm equivalent), creating rectilinear or cylindrical display images covering 120° x 60° at 267 megapixels per eye.

\* [ainsworth@qwerty.com](mailto:ainsworth@qwerty.com)

## 1.1 StarCAVE display

The StarCAVE is an ideal VR environment for displaying the capabilities of stereo panoramas created by the dual-camera system, as shown in Figure 1. The resolution of the image capture system is hundreds of megapixels per eye and easily matches the resolution of this display.



Figure 1. The person in this photo is viewing a spherical panorama created by the dual-camera imaging system. Full stereo view is presented in this StarCAVE virtual reality display at California Institute for Telecommunications and Information Technology<sup>[4]</sup>. Resolution of the image capture system is hundreds of megapixels per eye.

## 1.2 Spherical projection

Multiple images are photographed and combined in stitching software to create a spherical projection that captures the complete  $360^\circ \times 180^\circ$  view. This process is repeated for the left and right eye positions. The resulting rectangular image for either position is a cylinder that represents a complete sphere, as shown in Video 1.

The video link shows how this image can appear when projected onto a spherical surface. This view is useful in checking the continuity of the image capture as part of the stitching process, and before processing to match the VR environment. This is particularly helpful in checking continuity of the stitched image at both nadir and zenith.

The rectangular view shown below, or the spherical projection shown in the video link, only represents half of the stereo capture. Resolution is limited to what can be conveniently viewed on the computer screen.



Video 1. Taliesin Entrance. This entrance to Frank Lloyd Wright's residence at Taliesin East, near Spring Green, Wisconsin, is photographed in a  $360^\circ$  panorama. Used by permission. <http://dx.doi.org/10.1117/12.872521.1>

## 2. ACQUISITION METHOD

The same photographic image capture technique is used for creating spherical stereoscopic projections for VR environments and high-resolution rectilinear or cylindrical images for multi-tiled rectangular displays. In both applications, an array of images is photographed and then processed in stitching software to create several types of projections, depending on the application. Image resolution can be as high as desired.

Spherical images will be processed as equirectangular projections and mapped to the dimensions of the VR system. For rectangular displays, either rectilinear or cylindrical projections can be created in the dimensions required by the display. With spherical image capture, resolution at the surface of the sphere is a function of the focal length of the lenses used. Rectangular image resolution at the surface of the display is a function of both focal length and the angle of capture.

### 2.1 Equipment

Two cameras are programmed to create two sets of images corresponding to the view from each eye. The assembly shown in Figure 2 can be programmed to cover a complete spherical view or any section of a sphere.



Figure 2. The dual-camera system is fully automated and can capture any number of stereo images covering a complete sphere or rectangle. Both nadir and zenith can be included, with the exception of a small footprint below the tripod.

**Panasonic LUMIX® GF-1** cameras<sup>[5]</sup> were selected because they offer the resolution capability and flexibility of the best digital SLR systems, while maintaining a small profile. These cameras offer several essential features for this application, including 12.1-megapixel resolution and the possibility of being mounted side-by-side at the appropriate distance for stereo separation. In this mounting configuration, the interocular distance is 70mm. The mounting brackets are supplied by Berezin Stereo Photography Products<sup>[6]</sup>.

The **GigaPan EPIC Pro Robotic Controller**<sup>[7]</sup> is used to automate the capture process. This programmability allows repeat image capture and the option for HDR (high dynamic range) processing and other options. GigaPan is a collaborative project between Carnegie Mellon University and NASA Ames Intelligent Systems Division's Robotics Group<sup>[8]</sup>. This robotic system can be programmed to accommodate any number of images, matching the configuration of VR and other display technologies.

The **Ainsworth CC-1 Dual-Camera Controller**<sup>[9]</sup> was developed specifically for this application, and provides an interface between the robotic unit and the dual cameras. This unit is necessary because the Panasonic GF-1 camera exposure and focus functions are activated by varying current levels, rather than by simple switch closure. The controller accepts the triggering output from the GigaPan Epic Pro system and supplies synchronous current pulses to the cameras.

## 2.2 Panorama photography

Our method of panorama photography is based on creating a horizontal row of multiple images as the camera position is precisely rotated. The vertical dimension is created by additional horizontal rows at various elevations. When combined in a stitching program, these images form a seamless rectangle, including wide field images up to 360° and vertical dimensions that include both poles. If the camera positioning is done precisely, the resulting image is continuous cylinder with minimal distortion. A typical cylinder, mapped as a spherical projection, is shown below in Video 2.



Video 2. Taliesin Courtyard. This view of the courtyard at Frank Lloyd Wright's Taliesin East, near Spring Green, Wisconsin, is photographed in a 360° panorama. Used by permission. <http://dx.doi.org/10.1117/12.872521.2>

## 2.3 Stereo-panorama photography

Panorama stitching software is designed specifically to overcome minor misalignments when combining multiple images. This blending process distorts images slightly so that adjacent photographs connect in a seamless manner. This built-in correction capability of stitching programs offers an opportunity to create seamless stereo images from image capture techniques that incorporate small amounts of inherent distortion characteristics. This makes it possible to rotate a camera pair about a central axis, creating two overlapping panoramas that preserve accurate stereo separation.

The left photo in Figure 3 shows a camera rotation method for creating 2D panoramas. The single camera is rotated around the point of zero parallax, shown by the red dot, while capturing overlapping images. The photo on the right shows the technique for rotating dual cameras around a central point, again shown by the red dot. In this case, the pair is rotated about a common axis located midway between the left and right cameras and on a line connecting their zero parallax points. The sequence of images photographed by this technique form seamless panoramas in full stereo.



Figure 3. The left camera is rotated about the point of zero parallax to create a 2D image. The dual-camera stereo system on the right is rotated around a central point that is midway between the zero parallax points of the two lenses.

The technique for rotating the cameras about a common point in the horizontal direction may be applied vertically as well. This creates two concentric spherical images covering the full 360° x 180° field of view. The GigaPan robotic unit provides programmed camera motion capture, automating the process. The amount of inherent distortion created by our dual-camera system is a direct function of the camera offset, or the distance from the zero parallax point to the center of rotation. Limiting this offset to approximately 35mm or less also limits the distortion effects of this technique to what could be compensated for later in the stitching process, provided that objects are no closer than approximately five feet from the point of rotation. Offsetting each camera by this amount provides a consistent 70mm interocular distance at all points in the resulting sphere.

## 2.4 Software

The software converts the two image arrays created by the dual-camera system into two matching stereo views. The requirements are organizing the images into collections that match the image capture, adjusting individual exposure and possibly white balance if necessary, stitching the individual images to create a panorama, creating a projection that matches the intended display, and making final adjustments. Our choices for this project include Adobe Bridge® and Adobe Lightroom®, in Figure 4, and PTGui Pro® and Photo Shop® in Figure 5.

Adobe Bridge or similar software can be used for viewing a large number of images, selecting the correct sequence if several exposures were shot, and matching corresponding left and right pairs. Adobe Lightroom can be used to make individual adjustments in light level, white balance, and contrast. Arranging the size of the thumbnail images to correspond to the width of the panorama gives a realistic view.

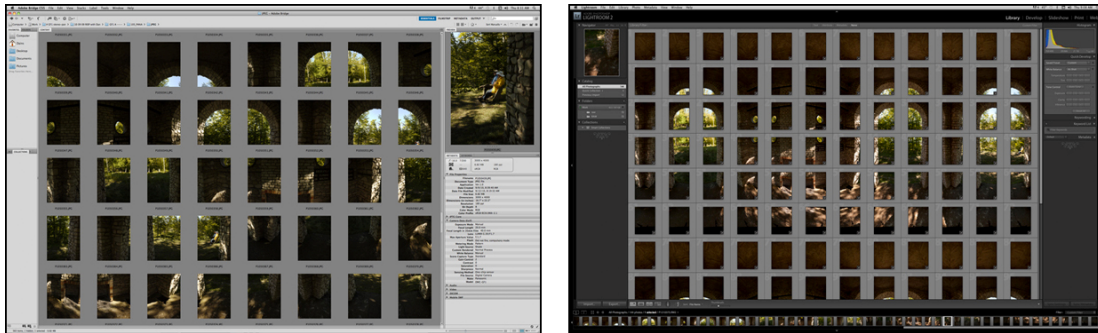


Figure 4. **Adobe Bridge** and **Adobe Lightroom** provide a workflow for sorting and adjusting images from the dual-camera system. This step organizes and adjusts images prior to the stitching process.

The PTGui software allows multiple copies to be operated in parallel. This feature permits easily switching back and forth between left and right views. After the initial alignment, the *Panorama Editor* is selected for each image. Alternating between the two views allows precise adjustment of the roll, pitch, and yaw variables via the *Numerical Transform Editor*. Vertical alignment of the two views is critical, and can be adjusted to within 0.1 degree with this method. Finally, the yaw parameter is adjusted to align the picture plane at the preferred distance from the camera. The time required for creating the image file depends on the resolution selected and the capability of the computer used. For a Mac Pro with 6GB RAM, processing times for a maximum resolution panorama can reach 12 hours or more.

Final adjustments for optimum contrast, color balance, saturation, sharpness, and similar parameters are done in Photo Shop. If both left and right images are combined in *Layers*, it is possible to switch rapidly back and forth between both views. This offers a quick and accurate test of stereo separation and balance.

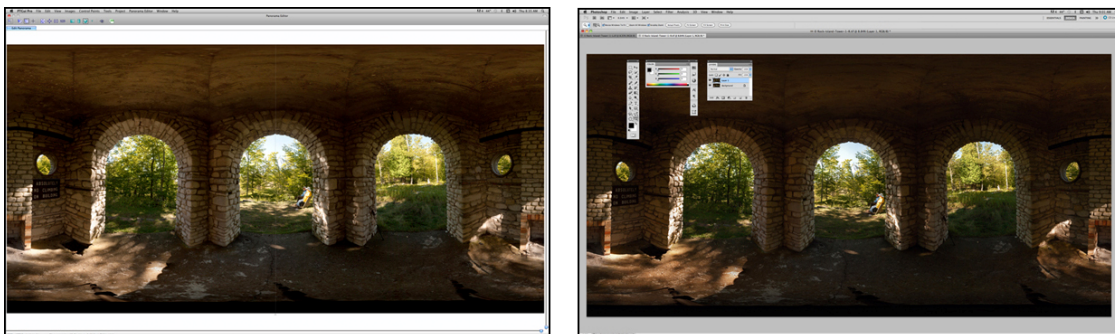


Figure 5. **PTGui Pro** stitching software is used to create the two matching stereo images and allows adjustment of roll, pitch, and yaw parameters. The final contrast, saturation, and sharpening are performed in **Photo Shop**.

### 3. SPHERICAL IMAGING

The same photographic image capture technique is used for creating both spherical and rectangular projections. Spherical images are processed as equirectangular and mapped to the dimensions of the VR environment. Rectangular images are processed as rectilinear or cylindrical images and mapped to the dimensions of the rectangular display.

#### 3.1 Photographic capture

A single rotation of the panorama head creates a horizontal row of images. Changing the vertical angle creates additional rows of images as shown below. In this example, fifteen images are required to cover the full 360° horizontal and five rows of images are required to cover the 180° vertical field. The five horizontal cylinders were shot at vertical angles of +70, +35, 0, and -35, -70 degrees. A set of images is created for each eye, separated by the interocular distance.

The composite image in Figure 6 shows all 75 individual photographs. This represents either a left or a right component of the stereo image. The entire stereo for the VR output is constructed from at least 150 individual photographs. If higher resolution is desired or if exposure bracketing is required, this number will rise dramatically.

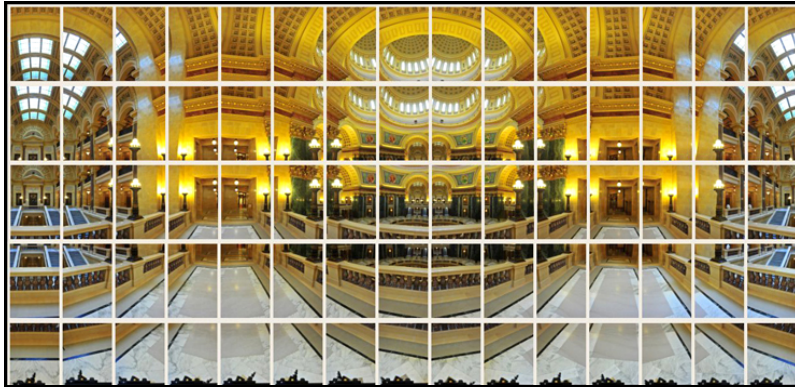


Figure 6. These 75 images will be combined to create a single left or right channel of the 360° by 180° stereo view.

#### 3.2 Equirectangular or cylindrical projection

Stitching software blends the separate images and creates a 360° by 180° projection, shown in Figure 7. This equirectangular or cylindrical projection will be subsequently mapped onto a corresponding sphere or cylinder in the VR environment.

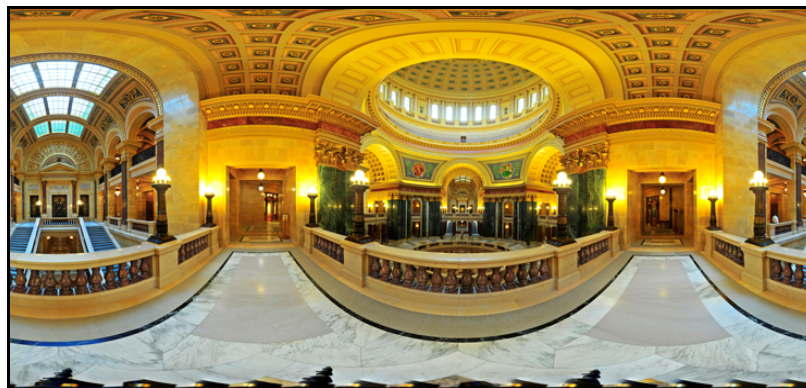


Figure 7. After processing in the stitching software, the separate left and right equirectangular projections. If comparison is made between the two images by switching back and forth between the two views, the stereo shift is apparent.

### 3.3 Spherical display -- QTVR

The Quick Time (QTVR) test shown in Video 3, if selected in the stitching software before the final projection is created, will map the image onto a sphere. The result is similar to the complete VR presentation, showing full rotation and both poles. This test does not show depth, but it does indicate whether normal perspective will be restored and all straight lines will appear normal when viewed later in the VR environment.



Video 3. Wisconsin Capitol. This test image shows how the equirectangular projection appears when mapped to a sphere, and how the final result will appear when viewed in the VR environment. <http://dx.doi.org/10.1117/12.872521.3>

### 3.4 VR display – NexCAVE

The two photographically created images are displayed on a cylinder or a sphere, depending on the projection used in the stitching. The radius from the viewer to the image is set at approximately infinity, or at the object of major interest. The two spherical images are displaced horizontally by the interocular distance, as seen from the viewer's perspective. Displacement equal to the interocular distance is maintained perpendicular to a plane aligned with the viewer's direction and perpendicular to a line between the eyes. This presents good stereo separation in whatever direction the viewer looks, even up and down. Objects at infinity will move with the viewer, contributing to an immersive experience. The NexCAVE in Figure 8 is a typical application.



Figure 8. The NexCAVE high-resolution VR display at King Abdullah University of Science and Technology<sup>[10]</sup> presents polarized images matching the left and right views.

## 4. RECTANGULAR IMAGING

Rectangular images are processed as rectilinear or cylindrical projections and may be used directly in high-resolution multi-tiled displays<sup>[11]</sup>, projection systems, and print. The correct visual perspective can be retained and further processing beyond conventional stitching software is not required.

### 4.1 Photographic capture

While spherical images for VR displays are frequently based on a 360° horizontal image capture, images for rectangular displays are limited to a 120° or less in the horizontal direction. Images wider than this cannot be rendered in rectilinear projection. The overall dimensions of the photographic capture will depend on the aspect ratio of the intended result. Resolution is a function of both the horizontal angle of the complete image and the focal length of the lenses used. This composite image in Figure 9 shows all 36 individual photographs, representing either a left or a right component of the stereo image. Shooting slightly beyond the desired frame allows for cropping later in the final images.

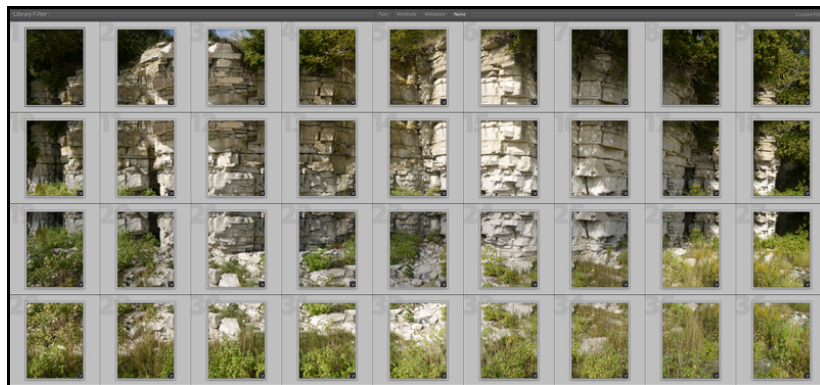


Figure 9. One channel of the original image capture is shown here as it appears in Adobe Lightroom2. The thumbnail size is adjusted so that the images in this view match the original order and layout in which they were shot. This arrangement facilitates making micro exposure adjustments to individual frames prior to stitching.

### 4.2 Rectilinear or cylindrical projection

The stitching software blends the separate images and creates a rectangular or cylindrical projection as shown in Figure 10. This image will be subsequently mapped onto a corresponding display. Image resolution is typically in the hundred-megapixel range, and can be as high as desired.



Figure 10. The stitching software produces separate left and right images in either cylindrical or rectilinear projection. These images may be combined for full stereo electronic display or large-scale anaglyph prints.



### 4.3 Rectangular display

Electronic displays in either 2D or 3D and large-scale anaglyph images can be produced by the same image capture technique. Shooting in full stereo and including wider angles of view than are strictly required allows for precise matching to the display. Both cylindrical and rectilinear projections are available in the stitching software and can be used to create accurate images in rectangular display format.

Even if the intended display is only 2D, as shown below in Figure 11, capturing the original scene so that it may also be shown in stereoscopic mode takes little additional capture time and is a reasonable investment for applications with future display technologies. Capture in 3D also allows for anaglyph display or printing.



Figure 11. A section of the original image is shown in this 40-megapixel display. The AESOP 40-tile display wall at King Abdullah University of Science and Technology was first prototyped at Calit2.

Renderings of reasonable architectural accuracy are possible with rectilinear projections, provided that the horizontal input angle is not too wide. The absolute maximum for rectilinear progressions is  $120^\circ$ . Several rectilinear sections can be created like this one from a single  $360^\circ$  image capture of an interior view. Architectural interiors like the one in Figure 12 can also be photographed and displayed accurately with this technique.



Figure 12. This architectural interior image captured in rectilinear progression and is  $110^\circ$  wide. Rectilinear projections like this show little distortion and may cover up to  $120^\circ$ .

## 5. VR DISPLAY METHOD

The stereo panorama images can be displayed in immersive virtual reality systems in a way that the user experiences being in the place the photographer was when the surround images were created. We use the COVISE or Electro visualization framework, which is compatible with a wide variety of immersive and non-immersive, 3D and 2D display environments, ranging from desktop computers to CAVEs. The key idea for stereo is to display the left and right eye images separately for each eye. That, along with the image being displayed at life size, will create a very realistic immersive experience.

The type of projection selected will depend on the overall angle of view of the intended display.

- For relatively small horizontal and vertical view angles, a rectilinear projection is textured onto a plane.
- For large horizontal angles and small vertical angles, a cylindrical projection is textured onto a cylinder.
- When both horizontal and vertical angles are large, an equirectangular projection is textured onto a sphere for the left and right eye respectively.

The photographic panoramas are produced such that the points at infinity are in the same place on the left and right eye image. This textured sphere then goes through the normal VR perspective projection for the left and right eye. If the sphere radius is large, the left and right eye projections on the screen are displaced by the eye positions. If the tracked viewer head is oriented to view the screen "square on", the two images on the screen are displaced by the interlocutor distance along a line parallel to the line between the eyes.

The correct disparity at infinity is produced by the VR projection. Disparities captured in the panorama closer than infinity are effectively subtracted from the disparity at infinity. We often set the radius of sphere at a point of interest in the scene. In an interior, we set the radius at the distance to the wall. This would move the disparity closer to the viewer, which is more appropriate for that kind of scene.

Since the direction of the disparity is parallel to the line between the eyes, there is minimum vertical disparity error as the viewer tilts her head. With a sphere of large radius, objects move with the viewer, which is correct for objects at infinity. The disparity at the nadir and zenith is approximately correct.

### 5.1 Advantages of this image capture and display method

The method described in this paper has a number of advantages. It is not dependent on the particular screen arrangements of the VR hardware systems. The stereo panoramas can be displayed without software or hardware modification on a large number of VR devices, including one-wall through six-wall CAVEs. Images created by this technique are also compatible with a variety of software systems, including COVISE<sup>[12]</sup>, CAVELib<sup>[13]</sup>, and Electro<sup>[14]</sup>.

If the VR software system has the ability to do different things for different eyes, a common property, then the method will work. We have ported the method to a large variety of VR systems including classic CAVE<sup>[15][16]</sup>, StarCAVE<sup>[1]</sup>, NexCAVE<sup>[2]</sup>, Varrier<sup>[17][18]</sup>, and CORNEA<sup>[3]</sup>.

There are limitations, however, in response to head motion. If the radius of the sphere is large, the entire scene moves with the viewer. This is correct for distant objects, but not for objects at a different radius than the sphere. There is also no "look around" and the viewpoint of the projection is from the camera position, not the viewer position.

### 5.2 COVISE visualization framework

In order to display the surround panorama images with COVISE, they need to undergo the following steps. The source data set is one large TIFF image file. A typical size would be 22k x 14k pixels, either with cylindrical or with spherical mapping. These images are expected to be RGB at 24 bit/pixel, but other formats could easily be supported as well.

In the first step we use ImageMagick to resize the image so that the number of pixels along each edge becomes a multiple of 4k. For example, the 22k x 14k image would get scaled to 16k x 8k. This size works well with the StarCAVE hardware, but it may be too much texture data for other systems. The resizing happens without maintaining the aspect ratio of the image; that is going to be corrected when the image is being displayed.

In the next step, we load the resized images (in case of 3D stereo images, there will be a left-eye and a right-eye image) into our COVISE plugin. Upon loading, our software will tile the image into 4k x 4k textures. The tessellation of the image will happen differently, depending on whether a cylindrical or spherical image is to be displayed.

### 5.3 Image conversion

For cylindrical images, every 4k texture tile get tessellated into 30 quads, each as high as the tile and  $1/30^{\text{th}}$  of the width of the tile. The number 30 is an empirical value, which we find, works well in the StarCAVE. Our COVISE plugin has a menu in which the user can select the radius of the cylinder that the image is to be displayed on. A good value for this radius we found to be 30 feet. From the radius, knowing with what vertical viewing angle the panorama was shot, we can calculate the height of the cylinder, and thus the height for each textured quad. The cylinder is positioned so that its center is in the user's head position, while its height is user adjustable; the user should set the height so that the horizon in the image is at the user's eye height.

In the case of a spherical image, our application will tessellate each 4k x 4k texture tile into 30x30 quads. Again, this is an empirical value which works well in the StarCAVE. The radius of the sphere is user adjustable, just like in the case of the cylinder; we use 30 feet by default, which is a good value for the StarCAVE. And again, the horizontal and vertical angle width of the panorama image is dependent on how the panorama was shot. Those angles and the radius fully define our sphere. The last remaining step is to position the sphere: we place its center in the user position. The height is adjustable, so that the user can move the horizon to eye height, which is where it would be in reality. The texture mapping for spherical images happens evenly across the tile. This means that towards the north and south pole of the sphere, the tiles shrink, down to a single point in the very poles. This assumes that the spherical panorama was shot all the way to the poles.

### 5.4 Configuration file

Every panorama image needs to come with information about its horizontal and vertical view angle, in order to allow for a correct display in VR. These and other parameters, including the user adjustable parameters in the menu can be configured or initialized in COVISE's configuration file. The configuration options for this entry:

- reye - right eye image
- leye - left eye image
- radius - world space radius of sphere/cylinder
- viewanglev - vertical view angle
- viewangleh - horizontal view angle
- camHeight - height offset for camera position
- segments - number of segments to use when rendering
- tsize - image tile size
- sphere - render as sphere instead of cylinder

Here is an example for a configuration file entry for a spherical stereo panorama image:

```
<Taliesin_Garden_13 reye="/state/partition1/data/SpherePans/Taliesin-Garden-13-R16_8.tif" leye="/state/partition1/data/SpherePans/Taliesin-Garden-13-L16_8.tif" radius="9144" viewanglev="180" viewangleh="360" camHeight="-1128.0" segments="30" tsize="4096" sphere="1" />
```

Our COVISE application offers the following configurable parameters for the user to change with dials: the height offset of the cylinder or sphere from the origin of the room coordinate system; the radius of the cylinder or sphere; the view angles the panorama covers; the number of tessellation segments per texture tile.

## 6. CONCLUSIONS

This system provides a compelling 3D experience with high-resolution photographic realism of architectural and natural environments. These stereo panoramas have been shown publicly to hundreds of viewers in a variety of VR display systems and VR software systems. Although the method does not provide viewer-centered perspective, it does utilize the stereoscopic and large angle of view properties of VR systems.

Two interesting developments have occurred as a result of this research. VR is expanded to now include photography as a medium for creating the entire immersive experience. Photography is expanded in scope as well, reaching beyond the conventions of the *frame* to include environmental imaging as created by the stereo panorama process.

Many of the adjustments, parameters, and concepts that are familiar in conventional photography need to be modified or even abandoned for photographic capture of VR environments. Other than selecting a location that's worth photographing in the first place, few skills learned from conventional photography remain intact. Many of the artistic considerations familiar to photographers are applicable in new ways.

### 6.1 Composition

Selecting a point of view is often the first consideration in creating a conventional photograph. With environmental imaging, however, there is no boundary to define the visual experience. This doesn't eliminate composition in the broad sense from consideration: it translates this aspect of the creative experience to encompassing more than two dimensions. The challenge becomes one of composing an image without reference to any edge. In this connection, selecting an environment for photographic capture has more in common with the ways we view sculpture or architecture than the artistic parameters associated with conventional photography.

### 6.2 Depth of field

Controlling the depth of field and region of focus is often a major creative element in conventional photography, rendering background or other information slightly softer and shifting point of view to a specific area of interest. This creative tool does not work the same way in VR because the viewer expects *everything* to always be in focus, just as it appears to be in the real world. If an object in virtual space does not immediately appear in perfect focus, the illusion is lost and virtual reality fails. With precise focus and aperture settings to preserve maximum depth of field for all visible objects we can present a virtual environment that mimics reality, regardless of the subject or portion of the image being scrutinized. To do this it is necessary to capture and visually recreate a scene the way we *think* we see, not as it actually appears to the eye.

### 6.3 Dynamic range

Capturing the complete dynamic range of a scene and recreating it in a print is the essence of the zone system<sup>[19]</sup> created by Ansel Adams<sup>[20]</sup> and others. Capturing dynamic range is challenging in conventional photography and even more problematic when capturing VR environments because of the wide range of light values. Exposure bracketing and HDR (high dynamic range) processing<sup>[21]</sup> is often the only solution. Fortunately, stitching software is designed to process HDR images automatically in both true HDR and image fusion modes. There are, however, drawbacks to this approach. With each additional set of bracketed images, processing time increases dramatically. This increased data collection taxes everything from camera buffers and storage media to post production times. Computer processing time can easily stretch to more than a day to calculate and display a single image, using current techniques and conventional computers.

The dynamic range of human vision is extensive, allowing people to easily adjust to a wide range of light values. Photographs, by comparison, are extremely limited. As a result, the same loss of detail in highlights and shadows that is typical of many conventional photographic prints will appear unreal in a virtual setting. This places additional demands

on the dynamic range that must be captured in order to provide a believable representation of what we would normally expect to see. With conventional photography, a major challenge involves adjusting the dynamic range of the image to match the very limited dynamic range of the photographic print. When digital displays are used, the restrictions of print no longer apply and dynamic range can be extended.

#### **6.4 Image resolution**

Resolution is unlimited in stereo panoramas. If higher resolution is required, longer focal lengths are used to provide more images for covering a given area or angle of view, directly increasing resolution in the final product. This would be good news to those who are convinced that more is better where megapixels are concerned, except that processing demands quickly limit image resolution to practical levels.

A reasonable approach would be to determine the maximum resolution the display media can accommodate and set the focal length and number of images required accordingly. Original image capture in higher resolution than is required in the current VR or other display is a reasonable choice, given that future display technology will probably continue to accommodate increased resolution capabilities. Similarly, photographing a scene in stereo where only a 2D output is required allows for future opportunities.

#### **6.5 Camera settings**

Automatic settings make modern digital cameras easy to use and give the photographer creative freedom to explore subject matter instead of being constantly immersed in myriad dials and details. Unfortunately, many of these settings must be disabled when shooting for VR capture. Focus, for example, can no longer be automatic. It's fine to use auto-focus to initially set the distance, but this feature must be turned off when shooting panoramas. Shifting focus during capture can result in major postproduction confusion.

Exposure should also remain constant. Variations in light level should be accommodated by bracketing all the images, not by changing exposure settings of individual images that will be combined in the final result. If the light level varies while a panorama is being photographed, processing the images in Adobe Lightroom or similar image correcting software can adjust for these variations prior to the stitching process. A more precise technique is to shoot the scene multiple times and create a correctly exposed result by hand-selecting individual images with corresponding light values.

Focal length and white balance must remain constant. The more advanced stitching software calculates the actual focal length used instead of relying on the EXIF focal length data encoded with the image file. If this parameter varies among images, the stitching software becomes completely befuddled. White balance must also be consistent for the complete scene. Allowing white balance to adjust automatically can result in wide variations that are extremely difficult to correct.

To create a believable VR experience based on a photographic representation, *all* of these critical camera settings must be optimized for that result. If any one parameter wavers significantly, the overall effect will be lost. The viewer will be simply immersed in "photographs" instead of the reality that's intended.

#### **6.7 Reality -- virtual and otherwise**

The mechanism of human visual perception abhors ambiguity and may go to extreme lengths to create reasonableness and continuity out of whatever visual information we perceive. This cognition is the primary contributor to our sense of what is real.

To create a virtual or alternate view of reality, sufficient information must be presented to make *that* view the preferred choice. When photorealism is added to the VR environment via environmental imaging, the data can be sufficiently compelling. It is no accident that we use the phrase "I see" when we actually mean comprehension and understanding. We live, after all, in a visually defined universe -- as opposed a dolphin's principally auditory environment and the canine's olfactory world. To us, seeing actually is believing -- even if it's *always* a virtual perception.

But our perceptions can be fooled, as with the *Necker Cube* illusion first examined by Swiss crystallographer Louis Albert Necker<sup>[22]</sup> in 1832 and shown here as Figure 13. There are several possible realities represented and, with insufficient data to definitively select one over the other, our actual perception of this event may waver. This is precisely the circumstance that allows us to perceive optical illusions of various sorts, and also allows us to intentionally manipulate the visual realities of others.

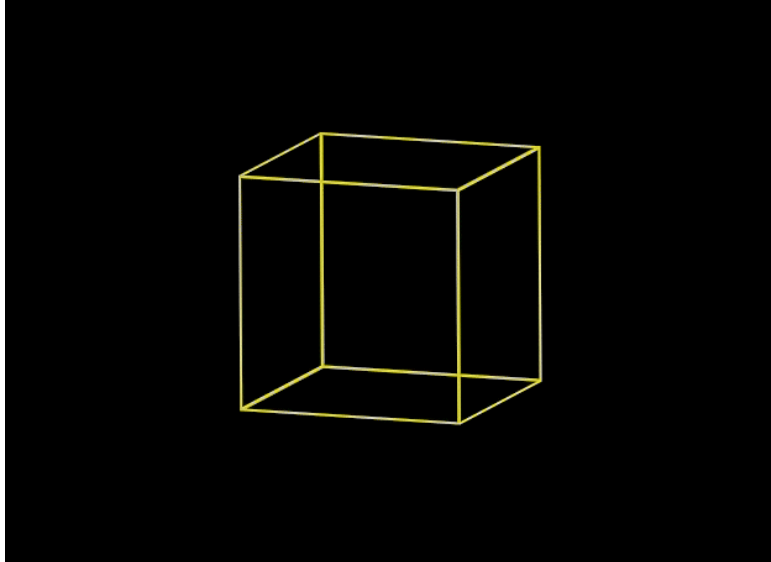


Figure 13. The Necker cube illustration shows how easily our visual perceptions may be altered when visual clues are ambiguous. An animated version is available here <http://www.qwerty.com/PIX/Necker.gif>

## ACKNOWLEDGEMENTS

This publication is based on work supported in part by Award No US-2008-107/SA-C0064, made by King Abdullah University of Science and Technology (KAUST).

UCSD received major funding from the State of California for the StarCAVE.

UIC receives major funding from the National Science Foundation (NSF), awards CNS-0420477 (Lambda Vision), OCI-0943559 (SAGE) and CNS-0959053 (NG-CAVE), and also receives support from Sharp Laboratories of America and the State of Illinois. Also, UCSD and UIC were lead institutions of the NSF funded OptIPuter project, award OCI-0225642.

CAVE™ is a trademark of the Board of Trustees of the University of Illinois.

*Taliesin Courtyard* © 2010 Dick Ainsworth and the Frank Lloyd Wright Foundation. Used by permission.

*Taliesin Entrance* © 2010 Dick Ainsworth and the Frank Lloyd Wright Foundation. Used by permission.

*Necker Cube in Rotation* -- Centre for Cognition; Donders Institute for Brain, Cognition, and Behavior

## REFERENCES

- [1] StarCAVE, [DeFanti08a] DeFanti, T., Dawe, G., Sandin, D., Schulze, J., Otto, P., Girado, J., Kuester, F., Smarr, L., and Rao, R., "The StarCAVE, a third-generation cave and virtual reality Optiportal," *Future Generation Computer Systems/The International Journal of Grid Computing: Theory, Methods and Applications*, Elsevier B.V., Vol 25, Issue 2, Nov. 2008, [http://ivl.calit2.net/wiki/index.php/StarCAVE\\_Measurement\\_Diagrams](http://ivl.calit2.net/wiki/index.php/StarCAVE_Measurement_Diagrams)
- [2] NexCAVE, USCD, <http://ucsdnews.ucsd.edu/video/nexcave.html>

- [3] CORNEA, KAUST, <http://www.kaust.edu.sa/research/labs/visualization.html>
- [4] California Institute for Telecommunications and Information Technology, (Calit2), <http://calit2.net>
- [5] LUMIX GF-1 Camera, Panasonic, <http://panasonic.net/avc/lumix/systemcamera/gms/gf1/>
- [6] Camera Mount, Berezin Stereo Camera Products, <http://www.berezin.com/3d>
- [7] GigaPan Epic Pro Robotic Controller, GigaPan Systems, <http://gigapansystems.com/gigapan-epic-pro.html>
- [8] Ames Intelligent Systems Division Robotics Group, (IRG), [ti.arc.nasa.gov/tech/asr/IntelligentRobotics/](http://ti.arc.nasa.gov/tech/asr/IntelligentRobotics/)
- [9] Ainsworth CC-1 Dual-Camera Controller, Ainsworth & Partners, <http://www.qwerty.com/CC1/>
- [10] King Abdullah University of Science and Technology, (KAUST), <http://www.kaust.edu.sa/>
- [11] High Resolution Multi-Tiled Displays, Calit2, <http://calit2.net/newsroom/article.php?id=1332>
- [12] COVISE, COLlaborative VIsualization and Simulation Environment, High Performance Computing Center Stuttgart (HLRS), <http://www.hlrs.de/organization/av/vis/covise>
- [13] CAVELib, Mechdyne Corporation,  
<http://www.mechdyne.com/IntegratedSolutions/software/products/CAVELib/CAVELib.htm>
- [14] Electro, Electronic Visualization Laboratory, <http://www.evl.uic.edu/rlk/electro/>
- [15] CAVE, T.A. DeFanti, D. Sandin, C. Cruz-Neira, A 'Room' with a 'View', IEEE Spectrum (1993) 30-33. 10/01/.
- [16] CAVE, Cruz-Neira, D. Sandin, T. DeFanti, R. Kenyon, J. Hart, The CAVE®: Audio-visual experience automatic virtual environment, Communications of the ACM (1992) 06/01/
- [17] VARRIER, Daniel J. Sandin, Todd Margolis, Jinghua Ge, Javier Girado, Tom Peterka, Thomas A. DeFanti, The Varrier autostereoscopic virtual reality display, ACM Transactions on Graphics: Proceedings of ACM SIGGRAPH 2005, Los Angeles, CA, July 31-August 4, 2005, Vol. 24, No. 3, pp. 894-903.
- [18] VARRIER, Cruz-Neira, D. Sandin, T. DeFanti, R. Kenyon, J. Hart, The CAVE®: Audio visual experience automatic virtual environment, Communications of the ACM (1992) 06/01/.
- [19] Lav, Brian. [Zone System: Step-by-Step Guide for Photographers], Buffalo, NY: Amherst Media, 2001
- [20] Adams, Ansel. [The Negative], New York Graphic Society, Reprinted, Boston: Little, Brown, & Company, 1995
- [21] High Dynamic Range (HDR) imaging, Wikipedia, [http://en.wikipedia.org/wiki/High\\_dynamic\\_range\\_imaging](http://en.wikipedia.org/wiki/High_dynamic_range_imaging)
- [22] Necker, Louis Albert, [http://en.wikipedia.org/wiki/Necker\\_Cube](http://en.wikipedia.org/wiki/Necker_Cube) .