Image-Based Rendering using Image Warping

Leonard McMillan
LCS Computer Graphics Group
Massachusetts Institute of Technology

In this report, I present a complete image-based rendering system. This includes the derivation of a mapping function from first principles, an algorithm for determining the visibility of these mapped points in the resulting image, and a method for reconstructing a continuous image from these mapped points. I refer to this type of mapping function as *image warping*, because it processes the elements of an image according to their image coordinates and produces outputs that are image coordinates in the resulting image.

In addition to the coordinates of the reference image additional information is required for each pixel. This information is related to the distance of the object seen at a particular pixel from the image plane. There are many different measures that can be used to describe this distance. Distance can be specified as *range values* describing the Euclidean distance from the visible object to image's center-of-projection. If the viewing or image plane is known and the coordinate system is chosen so that the normal of this plane lies a unit distance along the z-axis, then this distance information is called *depth* or the pixel's *z-value*. However, there are many other reasonable choices for representing this same distance. For instance distance values can be described indirectly by to the relative motion of image points induced by a change in the camera's position, this distance representation is frequently called *optical flow*, and it is inversely related to the point's range. *Disparity* and *projective-depth* are two more representations of distance for which a warping equation can be developed. The choice of a distance metric often depends on knowing some additional information about how the image was formed (ex. knowledge of the image plane), but in some applications knowledge of this relationship will be unnecessary to perform the desired image warp. Rather than selecting a particular distance measure, and deriving the warping function based on it, I will instead develop a notion of depth that leads to the simplest expression for the desired warping function.

The warping function developed here will not be a one-to-one mapping. In those places where multiple points map to the same result a method to resolve which of the candidate points is visible is required. I will describe a simple method for determining these visible regions. This method will not rely on any geometric information from the scene, but only on the change in pose between the reference and viewing positions.

Finally, since an image will usually be represented as two-dimensional array of discrete samples, reconstruction methods are developed so that the transformed discrete points of the reference image can be used to estimate the appearance of the desired continuous image. I will suggest two methods for this reconstruction.

1. From Images to Rays

A perspective image describes a collection of rays from a given viewing position. Relative to this position any particular ray is uniquely determined by two angles. While the use of two angles sufficiently

describes the full range and dimension of the set of rays, it is not a very convenient representation for analysis or computation. Here I will consider an alternative parameterization of rays based on a general *planar-pinhole camera model*. This is the same planar-pinhole camera model that is commonly used in traditional three-dimensional computer graphics and computer vision.

The planar-pinhole camera is an idealized device for describing the rays that pass through a single point in space, called the *center-of-projection*, and are contained within some solid angle defined by a bounded planar section, called the *image plane*. This solid angle is well defined as long as the center-of-projection does not lie on the extended image plane. As the bounds of the image plane are extended indefinitely, the solid angle approaches 2π steradians, exactly half of the visual sphere.

Consider the rays emanating from the origin of a three-dimensional system with basis vectors $(\hat{i}, \hat{j}, \hat{k})$. Suppose also, that a second two-dimensional coordinate system is defined in the image plane, allowing each point on it to be identified by an image coordinate, (u,v), where *u* and *v* are scalar multiples of the two basis vectors, $(\hat{s}, \hat{t})$, defined in the image plane. Points on this image-plane can be given coordinates that are specified relative to this coordinate system. These points, along with their assigned coordinates, are referred to as *image-space points*. Without loss of generality, we can assume that the origin of image space lies at one of the corners of the bounded image plane. The following figure depicts these coordinate systems:

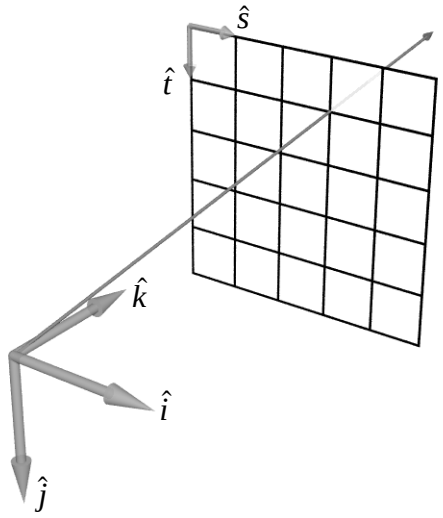

Figure 1: Mapping image-space point to rays

Each image-space point can be placed into one-to-one correspondence with a ray that originates from the Euclidean-space origin. This mapping function from image-space coordinates to rays can be described with a linear system:

$$\bar{d} = \begin{bmatrix} d_i \\ d_j \\ d_k \end{bmatrix} = \begin{bmatrix} a_i & b_i & c_i \\ a_j & b_j & c_j \\ a_k & b_k & c_k \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{P} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

Equation 1: Mapping from image coordinates to coordinates in three-space

where the coordinate of the image-space point is specified by the coordinate (u, v), and the resulting vector $\bar{d}$ represents the corresponding ray's direction. The entries of the mapping matrix, $\mathbf{P}$, can be easily understood by considering each column as a vector, $\bar{a}$, $\bar{b}$, and $\bar{c}$ in the same coordinate system as $\bar{d}$. This relationship is shown in Figure 2, where $\bar{a}$ and $\bar{b}$ are images of the $\hat{s}$ and $\hat{t}$ basis vectors in the $(\hat{i}, \hat{j}, \hat{k})$ coordinate system, and $\bar{c}$ is a vector from the ray origin to the origin of the image plane. Thus, while this parameterization is general, it still has a reasonable physical interpretation.

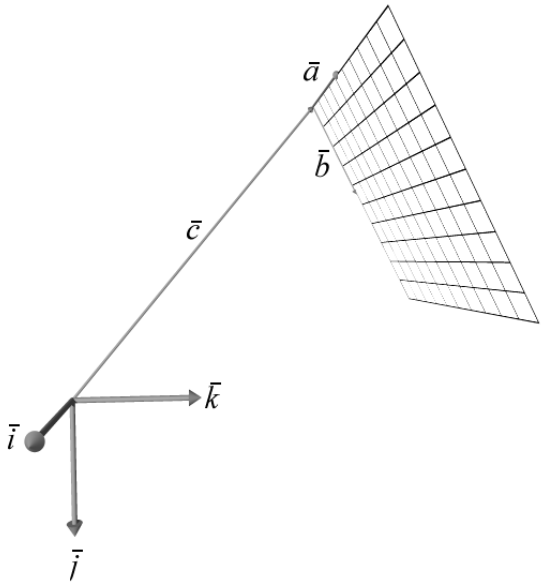

Figure 2: Relationship of the image-space basis vectors to the ray origin

The mapping function from image coordinates to rays is not uniquely defined. Any scalar multiple of the $\mathbf{P}$ matrix will also yield an equivalent set of image-space point-to-ray correspondences. This independence of scale is a consequence of the fact that the space of possible directions from any point in a three-dimensional space can be described using only two parameters (i.e., two angles). Thus, any representation of rays that uses unconstrained three-dimensional vectors will allow for multiple representations of the same ray.

2. A Warping Equation for Synthesizing Projections of a Scene

Equipped with only the simple planar-pinhole-camera model described in the previous section, an image-warping equation, which remaps those rays visible from a given viewpoint to any arbitrary viewing position, can be derived. As before, I will refer to the source image, or domain, of the warp as the *reference image*, and the resulting image, after the mapping is applied, as the *desired image*. For the moment, I will assume that both the pinhole-camera parameters of the desired and reference views are known. In addition, I will also assume that a single scalar value, called *generalized disparity* or *projective depth*, is known for

all points of the reference image. The precise nature of this quantity will be discussed in more detail later in this section. For the moment it is sufficient to say that this quantity can be determined from a set of correspondences specified between images.

Consider the implications of the same point, $\dot{X}$, being sighted along rays from two different centers-of-projection, $\dot{C}_1$ and $\dot{C}_2$, specified relative to their pinhole camera models. The following diagram illustrates this configuration.

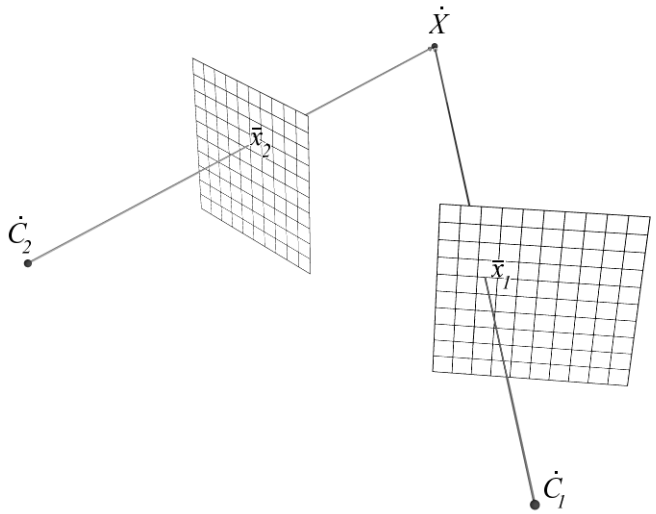

Figure 3: A point in three-dimensional space as seen from two pinhole cameras

The image coordinate, $\bar{x}_1$, in the first image determines a ray via the pinhole camera mapping $\bar{d}_1 = P_1\bar{x}_1$ with an origin of $\dot{C}_1$. Likewise, the image coordinate, $\bar{x}_2$, in the second image determines a ray, $\bar{d}_2 = P_2\bar{x}_2$, with origin $\dot{C}_2$. The coordinate of the point $\dot{X}$ can, therefore, be expressed using either of the following:

$$\dot{X} = \dot{C}_1 + t_1\mathbf{P}_1\bar{x}_1 = \dot{C}_2 + t_2\mathbf{P}_2\bar{x}_2$$

Equation 2: Specification of a 3D point in terms of pinhole-camera parameters

where t_1 and t_2 are the unknown scaling factors for the vector from the origin to the viewing plane which make it coincident with the point $\dot{X}$. This expression can be reorganized to give

$$\tfrac{t_2}{t_1}\mathbf{P}_2\bar{x}_2 = \tfrac{1}{t_1}\left(\dot{C}_1 - \dot{C}_2\right) + \mathbf{P}_1\bar{x}_1$$

Equation 3: Transformation of a ray in one camera to its corresponding ray in another

The left-hand side of this expression is now a ray, as is the second term on the right hand side. If we relax our definition of equivalence to mean “equal down to some non-zero scale factor” (which is consistent with the notion that rays having the same direction are equivalent regardless of the length of the three-space vector specifying this direction), then the t_2/t_1 factor can be eliminated. I will use the symbol,

$\doteq$, to represent this equivalence relationship. Alternatively, we could take advantage of the property that both $\mathbf{P}_1$ and $\mathbf{P}_2$ are defined independent of scale, to absorb the scalar quantity, t_2/t_1, into the matrix $\mathbf{P}_2$. Substituting the *generalized disparity* term $\delta(\bar{x})$ for $\frac{1}{t_1}$ gives

$$\mathbf{P}_2\bar{x}_2 \doteq \delta(\bar{x}_1)\left(\dot{C}_1 - \dot{C}_2\right) + \mathbf{P}_1\bar{x}_1$$

Equation 4: Simplified planar ray-to-ray mapping

The name, *generalized disparity*, comes from the notion of stereo disparity. In the normal depth-from-stereo case, the cameras are assumed to have a particular geometric configuration. Both image planes are required to have the same pinhole-camera model. The vector connecting the centers-of-projection must be parallel to both image planes. And, the coordinate system is selected so that the $\hat{i}$ basis vector of the camera space is parallel to the $\hat{s}$ basis vector of the image planes, as shown below.

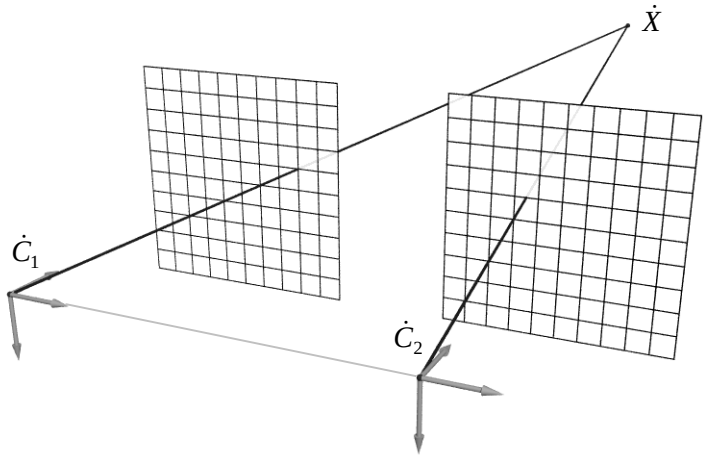

Figure 4: The depth-from-stereo camera configuration

This configuration can be accomplished either by accurate alignments of the cameras or by a post-processing rectification (using re-projection see Equation 13) of the acquired data. Under these conditions the planar ray-to-ray mapping equation simplifies to

$$\mathbf{P}_1\bar{x}_2 \doteq \delta(\bar{x}_1)\left(\begin{bmatrix} C_{1x} \\ C_{1y} \\ C_{1z} \end{bmatrix} - \begin{bmatrix} C_{2x} \\ C_{2y} \\ C_{2z} \end{bmatrix}\right) + \mathbf{P}_1\bar{x}_1 = \delta(\bar{x}_1)\begin{bmatrix} C_{1x} - C_{2x} \\ 0 \\ 0 \end{bmatrix} + \mathbf{P}_1\bar{x}_1$$

Equation 5:

Multiplying both sides by the inverse of the pixel-to-ray mapping gives

$$\bar{x}_2 \doteq \delta(\bar{x}_1)\mathbf{P}_1^{-1}\begin{bmatrix} C_{1x} - C_{2x} \\ 0 \\ 0 \end{bmatrix} + \bar{x}_1$$

Equation 6:

The alignment constraint on the camera-space and image-space basis vectors (that the $\hat{i}$ basis vector of the camera space is parallel to the $\hat{s}$ basis vector) implies that a unit step in image space produces a unit step in camera space. This is equivalent to

$$\mathbf{P}_1 = \begin{bmatrix} 1 & p_{12} & p_{13} \\ 0 & p_{22} & p_{23} \\ 0 & p_{23} & p_{33} \end{bmatrix} \quad \text{so} \quad \mathbf{P}_1^{-1} = \begin{bmatrix} 1 & q_{12} & q_{13} \\ 0 & q_{22} & q_{23} \\ 0 & q_{23} & q_{33} \end{bmatrix}$$

Equation 7:

After multiplying through and reorganizing terms we get the following:

$$\bar{x}_2 - \bar{x}_1 = \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} - \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = \delta(\bar{x}_1) \begin{bmatrix} C_{1x} - C_{2x} \\ 0 \\ 0 \end{bmatrix}$$

Equation 8:

Thus, when camera geometries satisfy the depth-from-stereo requirement, the image-space coordinates of corresponding points differ only along a single dimension in the images. The size of this change is proportional to both the distance between the centers-of-projection, which is often called the stereo baseline, and the generalized disparity quantity, which has units of pixels per unit length. Therefore, generalized disparity, $\delta(\bar{x})$, is equivalent to stereo disparity, $u_2 - u_1$, when normalized by dividing through by the length of the baseline.

$$\delta_{stereo}(\bar{x}) = \frac{u_2 - u_1}{C_{1x} - C_{2x}}$$

Equation 9:

The term, *projective depth*, is sometimes used instead of generalized disparity. This name is somewhat misleading since it increases in value for points closer to the center-of-projection. Generalized disparity, $\delta(\bar{x}_1)$, is inversely related to depth by a constant scale factor, and to range by a scale factor that varies from point to point on the viewing plane.

Let us consider further the generalized-disparity term, $\delta(\bar{x}_1)$. We can construct the following diagram of Equation 4 in the plane containing the three points $\dot{C}_1$, $\dot{C}_2$, and a visible point $\dot{X}$.

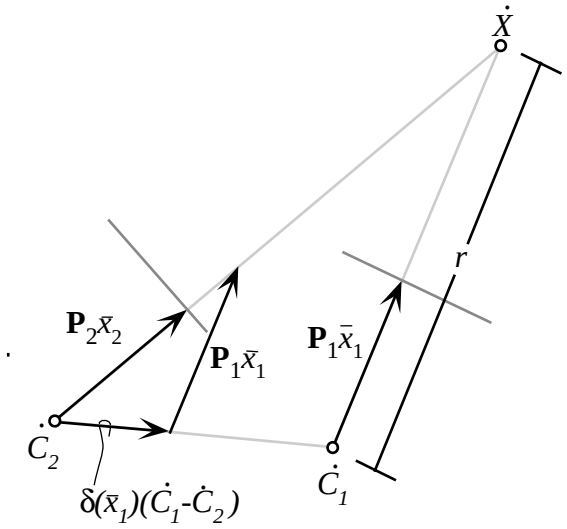

Figure 5: A point as seen from two images

Using similar triangles it can be shown that

$$\frac{r}{\left|\dot{C}_1 - \dot{C}_2\right|} = \frac{\left|\mathbf{P}_1\bar{x}_1\right|}{\delta(\bar{x}_1)\left|\dot{C}_1 - \dot{C}_2\right|}$$

Solving for generalized disparity gives the following expression:

$$\delta(\bar{x}_1) = \frac{\left|\mathbf{P}_1\bar{x}_1\right|}{r}$$

Thus, the generalized disparity depends only on the distance from the center-of-projection to the position on the image plane where the point is observed and the range value of the actual point. The image-space coordinate in the second image of the actual point can be found using the following transformation:

$$\bar{x}_2 \doteq \delta(\bar{x}_1)\mathbf{P}_2^{-1}\left(\dot{C}_1 - \dot{C}_2\right) + \mathbf{P}_2^{-1}\mathbf{P}_1\bar{x}_1$$

Equation 10: Planar image-warping equation

Since the generalized-disparity term, $\delta(\bar{x}_1)$, is independent of both the desired center-of-projection, $\dot{C}_2$, and the desired pinhole viewing parameters, $\mathbf{P}_2$, Equation 10 can also be used to determine the image-space coordinate of the observed point on any other viewing plane. This is accomplished by simply substituting the desired center-of-projection and pinhole-camera model into Equation 10.

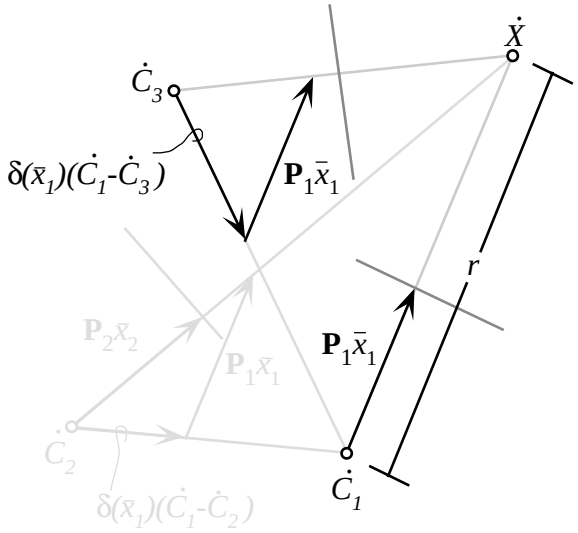

Figure 6: A third view of the point $\dot{\mathbf{X}}$

Figure 6 illustrates how the planar warping equation (Equation 10) can be used to synthesize arbitrary views. The third viewing position, $\dot{C}_3$, need not be in the same plane as $\dot{C}_1$, $\dot{C}_2$ and $\dot{X}$. Figure 6 also depicts how generalized disparity is an invariant fraction of the baseline vector. This fraction of the baseline vector applies to all potential views, and therefore, it can be used to align corresponding rays of a given reference image to their direction in any desired image. In addition, the resulting projection will be consistent to a fixed three-dimensional point. Generalized disparity is a scalar quantity that determines how a translation of the center-of-projection affects the coordinates of points in image space. The remaining matrix quantity, $\mathbf{P}_2^{-1}\mathbf{P}_1$, determines how changes in the pinhole-camera model, independent of translation, affect the coordinates of points in image space. A further explanation of this last claim will be presented in the next section.

The warping equation can be used to synthesize arbitrary views of a given reference image via the following procedure. Given a reference image, the matrix describing its planar-pinhole projection model, $\mathbf{P}_1$, its center-of-projection, $\dot{C}_1$, a generalized-disparity value, $\delta(\bar{x}_1)$, for each pixel, and the center-of-projection of the desired image, $\dot{C}_2$, and its projection model, the mapping of the reference image to a desired image can be computed. Using the vectors described in the generalized pinhole-camera model, the warping equation can be rewritten as

$$\alpha\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} \bar{a}_2 & \bar{b}_2 & \bar{c}_2 \end{bmatrix}^{-1}\left(\dot{C}_1 - \dot{C}_2\right)\delta(u_1, v_1) + \begin{bmatrix} \bar{a}_2 & \bar{b}_2 & \bar{c}_2 \end{bmatrix}^{-1}\begin{bmatrix} \bar{a}_1 & \bar{b}_1 & \bar{c}_1 \end{bmatrix}\begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix}$$

where α is an arbitrary scale factor. This matrix sum can be rewritten as shown below:

$$\alpha\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} \bar{a}_2 & \bar{b}_2 & \bar{c}_2 \end{bmatrix}^{-1}\begin{bmatrix} \bar{a}_1 & \bar{b}_1 & \bar{c}_1 & \left(\dot{C}_1 - \dot{C}_2\right) \end{bmatrix}\begin{bmatrix} u_1 \\ v_1 \\ 1 \\ \delta(u_1, v_1) \end{bmatrix}$$

Multiplying both sides by the determinant of $\mathbf{P}_2$ and substituting for its inverse gives the following 4×3 matrix equation:

$$\alpha\left(\bar{c}_2 \cdot \left(\bar{a}_2 \times \bar{b}_2\right)\right)\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} \bar{a}_1 \cdot \left(\bar{b}_2 \times \bar{c}_2\right) & \bar{b}_1 \cdot \left(\bar{b}_2 \times \bar{c}_2\right) & \bar{c}_1 \cdot \left(\bar{b}_2 \times \bar{c}_2\right) & \left(\dot{C}_1 - \dot{C}_2\right) \cdot \left(\bar{b}_2 \times \bar{c}_2\right) \\ \bar{a}_1 \cdot \left(\bar{c}_2 \times \bar{a}_2\right) & \bar{b}_1 \cdot \left(\bar{c}_2 \times \bar{a}_2\right) & \bar{c}_1 \cdot \left(\bar{c}_2 \times \bar{a}_2\right) & \left(\dot{C}_1 - \dot{C}_2\right) \cdot \left(\bar{c}_2 \times \bar{a}_2\right) \\ \bar{a}_1 \cdot \left(\bar{a}_2 \times \bar{b}_2\right) & \bar{b}_1 \cdot \left(\bar{a}_2 \times \bar{b}_2\right) & \bar{c}_1 \cdot \left(\bar{a}_2 \times \bar{b}_2\right) & \left(\dot{C}_1 - \dot{C}_2\right) \cdot \left(\bar{a}_2 \times \bar{b}_2\right) \end{bmatrix}\begin{bmatrix} u_1 \\ v_1 \\ 1 \\ \delta(u_1, v_1) \end{bmatrix}$$

Equation 11: 4×3 matrix formulation of warping equation

This results in the following rational expressions for computing the reprojection of pixel coordinates from a reference image to the coordinates of a desired image:

$$u_2 = \frac{\bar{a}_1 \cdot \left(\bar{b}_2 \times \bar{c}_2\right)u_1 + \bar{b}_1 \cdot \left(\bar{b}_2 \times \bar{c}_2\right)v_1 + \bar{c}_1 \cdot \left(\bar{b}_2 \times \bar{c}_2\right) + \left(\dot{C}_1 - \dot{C}_2\right) \cdot \left(\bar{b}_2 \times \bar{c}_2\right)\delta(u_1, v_1)}{\bar{a}_1 \cdot \left(\bar{a}_2 \times \bar{b}_2\right)u_1 + \bar{b}_1 \cdot \left(\bar{a}_2 \times \bar{b}_2\right)v_1 + \bar{c}_1 \cdot \left(\bar{a}_2 \times \bar{b}_2\right) + \left(\dot{C}_1 - \dot{C}_2\right) \cdot \left(\bar{a}_2 \times \bar{b}_2\right)\delta(u_1, v_1)}$$

$$v_2 = \frac{\bar{a}_1 \cdot \left(\bar{c}_2 \times \bar{a}_2\right)u_1 + \bar{b}_1 \cdot \left(\bar{c}_2 \times \bar{a}_2\right)v_1 + \bar{c}_1 \cdot \left(\bar{c}_2 \times \bar{a}_2\right) + \left(\dot{C}_1 - \dot{C}_2\right) \cdot \left(\bar{c}_2 \times \bar{a}_2\right)\delta(u_1, v_1)}{\bar{a}_1 \cdot \left(\bar{a}_2 \times \bar{b}_2\right)u_1 + \bar{b}_1 \cdot \left(\bar{a}_2 \times \bar{b}_2\right)v_1 + \bar{c}_1 \cdot \left(\bar{a}_2 \times \bar{b}_2\right) + \left(\dot{C}_1 - \dot{C}_2\right) \cdot \left(\bar{a}_2 \times \bar{b}_2\right)\delta(u_1, v_1)}$$

Since the centers-of-projection and the planar-pinhole camera models for the reference and desired images are fixed for a given mapping, the warping equation simplifies to a pair of constant-coefficient linear rational expressions of the form

$$r(u_1, v_1, \delta(u_1, v_1)) = w_{11}u_1 + w_{12}v_1 + w_{13} + w_{14}\delta(u_1, v_1)$$

$$s(u_1, v_1, \delta(u_1, v_1)) = w_{21}u_1 + w_{21}v_1 + w_{23} + w_{24}\delta(u_1, v_1)$$

$$t(u_1, v_1, \delta(u_1, v_1)) = w_{31}u_1 + w_{32}v_1 + w_{33} + w_{34}\delta(u_1, v_1)$$

$$u_2 = \frac{r(u_1, v_1, \delta(u_1, v_1))}{t(u_1, v_1, \delta(u_1, v_1))}$$

$$v_2 = \frac{s(u_1, v_1, \delta(u_1, v_1))}{t(u_1, v_1, \delta(u_1, v_1))}$$

Equation 12: Warping equation as rational expressions

The mapping of a point in the reference image to its corresponding point in the desired image can be computed using nine adds, eleven multiplies, and one inverse calculation. When points of the reference image are processed sequentially, the number of adds is reduced to six, and the number of multiplies is reduced to five. Additional tests for a positive denominator, $t(u_1, v_1, \delta(u_1, v_1))$, and a valid range of the numerator can avoid two multiplies and the inverse calculation. This operation is the equivalent of screen-space clipping in traditional three-dimensional computer graphics.

3. Relation to Previous Results

Many other researchers [Szeliski96] [Faugeras92] have described similar warping equations. In most of these applications the image-space coordinates of the points $\bar{x}_1$ and $\bar{x}_2$ were given, and the projective depth, $\delta(\bar{x})$, was the quantity to be solved for. When this equation is used for image warping, coordinates of image points in the desired view, $\bar{x}_2$, are computed from points in the reference image, $\bar{x}_1$, and their projective depths.

This warping equation is also closely related to several other well-known results from computer graphics, image-warping, and projective geometry. Consider the situation where the reference image and the desired view share a common center-of-projection. In this case the planar-warping equation simplifies to

$$\bar{x}_2 \doteq \mathbf{P}_2^{-1}\mathbf{P}_1\bar{x}_1$$

Equation 13: Image reprojection

This illustrates the well known result that images defined on planar viewing surfaces sharing a common center-of-projection are related by a *projective transformation* or *planar homography*.

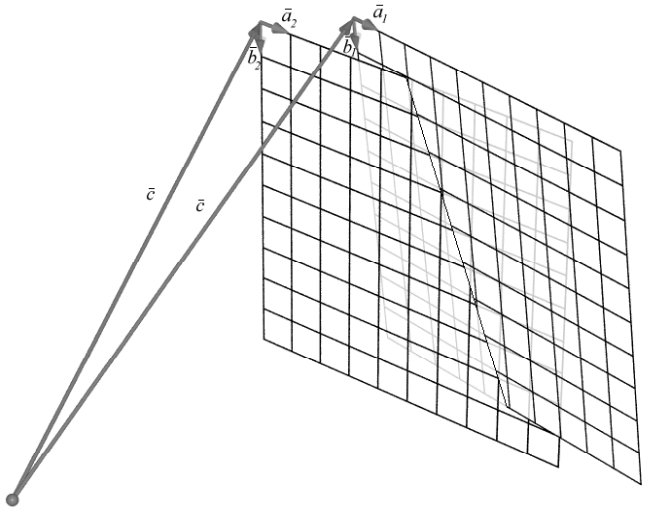

Figure 7: Reprojection of an image with the same center-of-projection

This projective transformation is merely the composition of the reference image's viewing matrix with the inverse of the desired image's viewing matrix, $\mathbf{H}_{reproject} = \mathbf{P}_2^{-1}\mathbf{P}_1$. This is indicative of the fact that, ignoring the clipping that occurs at the boundaries of the view plane, a change of viewing surface does not change the set of rays visible from the center-of-projection. It only changes their spatial distribution on the viewing surface. I will refer to mappings of this sort as reprojections.

A second well known result from the fields of computer graphics and projective geometry is that all images of a common planar surface seen in planar projection are also related by a projective transform as long as the plane does not project to a line. This result is the underlying basis for the texture mapping of images onto planar surfaces. It allows for the rasterization of textured planar primitives in screen space using a rational linear expression (an alternate formulation for a projective transform). The figure below illustrates the projection of a planar region onto several viewing planes and the resulting image.

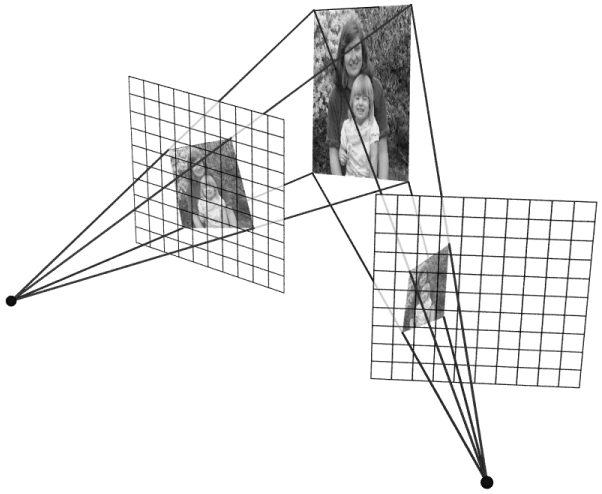

Figure 8: A planar region seen from multiple viewpoints

The mapping function that describes the possible views of a three-dimensional planar surface can be derived as a special case of Equation 10 by the following progression. The equation of a plane in the coordinate system having the reference image's center-of-projection as its origin is given by

$$\begin{bmatrix} A & B & C \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = D$$

Equation 14: Equation of a plane

where the scalars *A*, *B*, *C*, and *D* are four parameters defining the plane, and x, y, and z are the coordinates of a point in space. These three-space coordinates can be re-expressed in terms of image coordinates by using a planar-pinhole model image-to-ray mapping function as follows:

$$\begin{bmatrix} A & B & C \end{bmatrix} t(u,v) \mathbf{P}_1 \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = D$$

where $t(u,v)$ is a multiple of the distance from the center-of-projection to the viewing plane for the ray. Dividing both sides by the scalar quantities appearing on opposite sides gives

$$\delta(u,v) = \frac{1}{t(u,v)} = \begin{bmatrix} \frac{A}{D} & \frac{B}{D} & \frac{C}{D} \end{bmatrix} \mathbf{P}_1 \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

The inverse of $t(u,v)$ is equivalent to the generalized-disparity term discussed in Equation 3. When the generalized-disparity value from above is substituted into the warping equation (Equation 10), the following expression results:

$$\bar{x}_2 \doteq \mathbf{P}_2^{-1}\left(\dot{C}_1 - \dot{C}_2\right)\begin{bmatrix} \frac{A}{D} & \frac{B}{D} & \frac{C}{D} \end{bmatrix} \mathbf{P}_1 \bar{x}_1 + \mathbf{P}_2^{-1} \mathbf{P}_1 \bar{x}_1 = \mathbf{P}_2^{-1}\left(\left(\dot{C}_1 - \dot{C}_2\right)\begin{bmatrix} \frac{A}{D} & \frac{B}{D} & \frac{C}{D} \end{bmatrix} + I\right) \mathbf{P}_1 \bar{x}_1$$

Equation 15: Mapping of a common plane seen at two centers-of-projection

The planar homography, $\mathbf{H}_{plane} = \mathbf{P}_2^{-1}\left(\left(C_1 - C_2\right)\begin{bmatrix} \frac{A}{D} & \frac{B}{D} & \frac{C}{D} \end{bmatrix} + I\right)\mathbf{P}_1$, is a projective mapping of the reference-image points on the plane to their coordinates in the desired image. When the reference and desired images share a common center-of-projection, the projective mapping reduces to the reprojection given in Equation 13 as expected.

The image plane of a reference image is but a plane in three-dimensional space. Therefore, its image in any reprojection is related by a projective mapping. The equation of the three-dimensional image plane of a given pinhole-camera model is given by

$$\left(\bar{a}_1 \times \bar{b}_1\right)^T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \bar{c}_1 \cdot \left(\bar{a}_1 \times \bar{b}_1\right)$$

Substitution into Equation 15, gives

$$\bar{x}_2 \doteq \mathbf{P}_2^{-1}\left(\left(\dot{C}_1 - \dot{C}_2\right)\frac{1}{\bar{c}_1 \cdot \left(\bar{a}_1 \times \bar{b}_1\right)}\left(\bar{a}_1 \times \bar{b}_1\right)^T + I\right)\mathbf{P}_1 \bar{x}_1 = \mathbf{P}_2^{-1}\left(\left(\dot{C}_1 - \dot{C}_2\right)\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} + \mathbf{P}_1\right)\bar{x}_1$$

which simplifies to

$$\bar{x}_2 \doteq \left(\begin{bmatrix} \bar{0} & \bar{0} & \mathbf{P}_2^{-1}\left(\dot{C}_1 - \dot{C}_2\right) \end{bmatrix} + \mathbf{P}_2^{-1}\mathbf{P}_1\right)\bar{x}_1$$

Equation 16: Projection of the reference image plane in the desired image

Notice that the projective mapping, $H_{viewplane} = \lfloor 0 \quad 0 \quad \mathbf{P}_2^{-1}(C_1 - C_2) \rfloor + \mathbf{P}_2^{-1}\mathbf{P}_1$, describes the projection of the reference image's viewing plane in the desired image.

4. Resolving visibility

The mapping function described by the planar image warping equation (Equation 10) is not one-to-one. The locus of potential points in three-space, $\dot{X}(t)$, that project to the same image coordinate, $\bar{x} = (u, v)$, is described by the center-of-projection and a planar pinhole-camera model using the following equation:

$$\dot{X}(t) = \dot{C} + t\mathbf{P}\bar{x}$$

Equation 17:

The parameter *t* identifies specific three-dimensional points that project to a given image-space coordinate. A policy of selecting the smallest of all positive *t* values for a given screen-space point can be used to determine visibility for opaque objects. This candidate *t* value is analogous to the z values stored in a z-buffer [Catmull74], or the ray-length maintained by a ray-casting algorithm [Appel68]. While the *t* parameter is an essential element of the reprojection process, (via its relationship to $\delta(\bar{x}_1)$) it is, surprisingly, not required to establish the visibility of a warped image.

In this section, an algorithm is presented for computing the visible surface at each image-space point of a desired image as it is being derived from a reference image via a warping equation. This is accomplished by establishing a warping order that guarantees a correct visibility solution. This ordering is established independently of the image contents. Only the centers-of-projection of the desired and reference images, as well as the pinhole-camera model for the reference image are needed.

In order to simplify the following discussion, the reference image is assumed to be stored as a two-dimensional array whose entries represent uniformly spaced image samples. This simplification is not strictly necessary for the algorithm to operate correctly, but it allows for a concise statement of the algorithm, and it is representative of many typical applications.

The approach of this visibility algorithm is to specify an ordering, or enumeration, of points from the reference image which guarantees that any scene element that is hidden by some other scene element in the desired image will always be drawn prior to its eventual occluder. This type of ordering is commonly called back-to-front. It is well known that a simple painter's algorithm [Rogers85] can be used to display any collection of scene elements with correct visibility when a back-to-front ordering can be established.

Given the reference image's center-of-projection, $\dot{C}_1$, and ray-mapping function, $\mathbf{P}_1$, and the desired center-of-projection, $\dot{C}_2$, the projection of $\dot{C}_2$ onto the reference image is first computed as follows:

$$\begin{bmatrix} e_x \\ e_y \\ e_z \end{bmatrix} = \mathbf{P}_1^{-1}\left(\dot{C}_2 - \dot{C}_1\right)$$

Equation 18:

An example of this projection is illustrated in Figure 9.

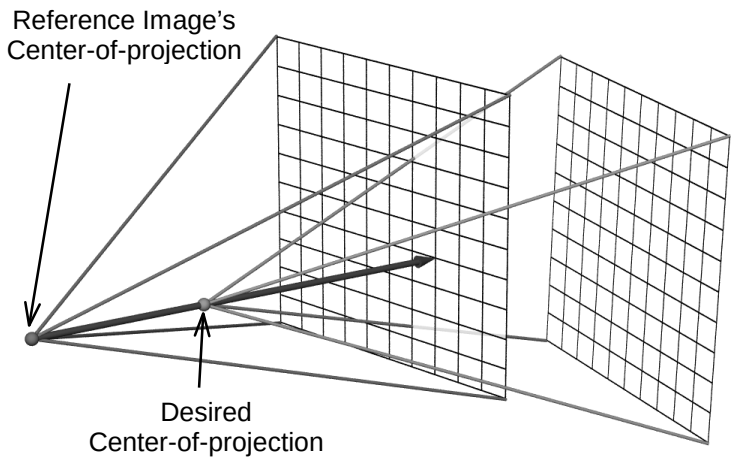

Figure 9: A desired center-of-projection projected onto the reference image

The image coordinate of $\dot{C}_2$ on the reference image is given by $\bar{e} = (e_x/e_z, e_y/e_z)$. It will fall into one of nine regions relative to the reference image shown below:

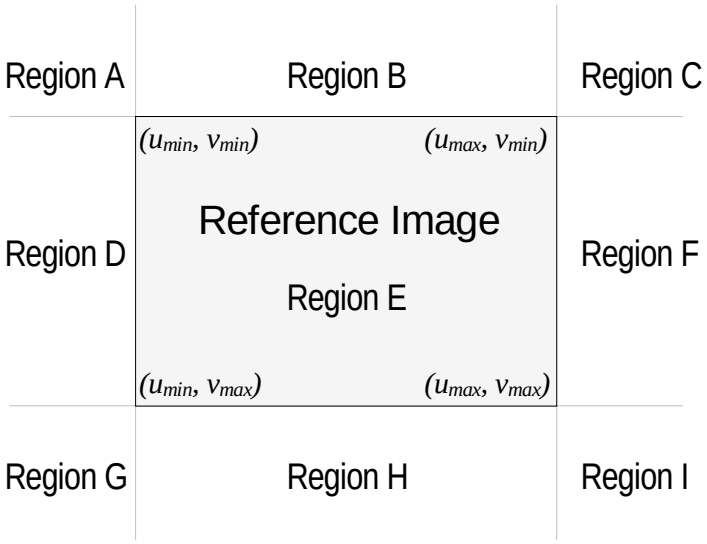

Figure 10: Figure of nine regions

Next, the reference image is subdivided into sheets that can be processed independently. Any set of orthogonal image-space basis vectors can be used to partition each sheet, but it is simplest to choose a coordinate basis aligned with the image's sampling grid. The reference image is partitioned into 1, 2, or 4 sheets depending on the image-space coordinates of the projected point, $\overline{e}$. When $\overline{e}$ projects within the domain of the reference image, the image is divided into four sections separated along the row and column of $\overline{e}$.

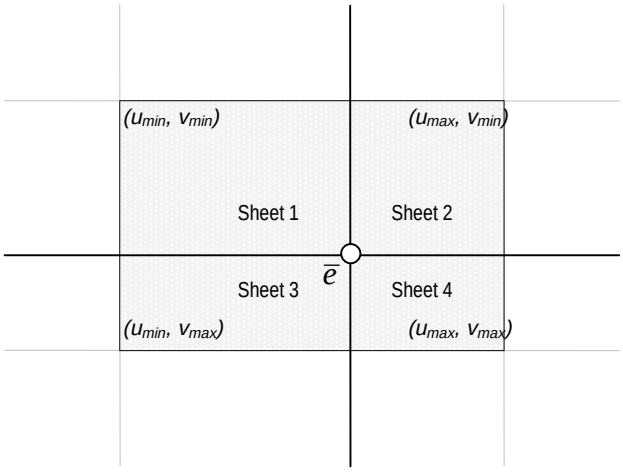

Figure 11: A desired center-of-projection that divides the reference image into 4 sheets

When only one coordinate of $\overline{e}$ falls within the reference image, (i.e., when $\overline{e}$ falls into regions B, D, F, and H) the image is subdivided into two sheets whose boundary is determined by either the row or column of the one coordinate that lies within the domain.

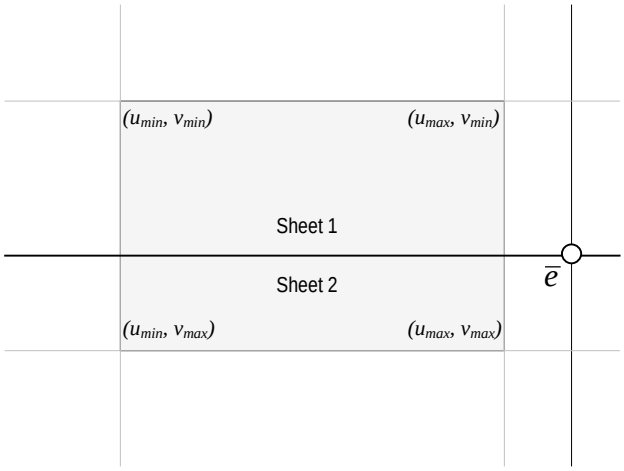

Figure 12: A desired center-of-projection that divides the reference image into 2 sheets

If neither component of $\bar{e}$ falls within the reference image's domain, (when $\bar{e}$ projects into regions A, C, H or J) then the entire image is treated as a single sheet.

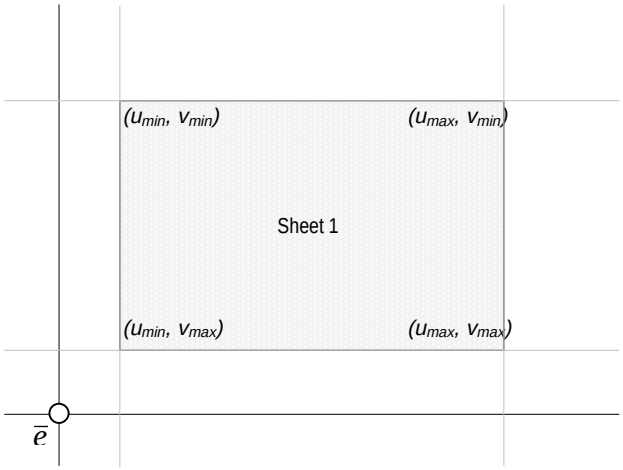

Figure 13: A desired center-of-projection that divides the reference image into 1 sheet

Once the reference image's domain is subdivided according to the projected position of the desired center-of-projection, the warping order for each sheet can be determined as follows. The sign of the e_z component from Equation 18 determines the enumeration direction. When e_z is non-negative the warping order of each sheet progresses toward the point $\bar{e}$, otherwise the warping progresses away from $\bar{e}$, as shown in the figure below. The case where e_z is zero indicates that the desired viewing position has no proper projection onto the reference image, because the vector $\dot{C}_1 - \dot{C}_2$ is parallel to the reference image's viewing plane. In this case, only one sheet will be enumerated, and the warping order progresses in the direction determined by the quadrant indicated by the signs of the remaining two vector components, e_x and e_y.

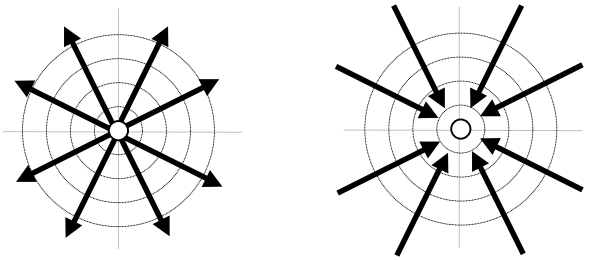

Figure 14: Enumeration direction

During the warp, each radial line originating from the projected image of the desired center-of-projection can be traversed independently. Alternatively, the warp can progress along either rows or columns of the sheets so long as the image of the desired center-of-projection, $\overline{e}$, is drawn at the appropriate time (i.e., p is drawn first when e_z is negative, and last otherwise), allowing either a row major or column major traversal. The advantage of the latter approach is that it allows the reference image's traversal to take maximum advantage of the access coherence of most memory systems.

The entire visibility algorithm involves three simple steps. First, the three-dimensional coordinate of the desired center-of-projection, $\dot{C}_2$, is projected on the reference image's viewing plane. Second, the image-plane is divided into sheets determined by the image-space coordinate of the projected center-of-projection, $\overline{e}$, and whose boundaries are aligned with the image-space coordinate axes[1]. Computing this coordinate involves a projective normalization. Finally, the sign of the planar normalizing element of the projective coordinate determines the traversal of the reference image[2].

The algorithm presented here is similar to Anderson's algorithm for bivariate functions [Anderson82]. The difference is that his visibility algorithm was defined for a different class of surfaces (i.e., a height field, or Monge patch, rather than a projective surface), and his algorithm enumerates the facets in a front-to-back occlusion order. Anderson's choice of a front-to-back order requires that some representation of the grid perimeter be maintained to aid in deciding what parts or edges of subsequent facets need to be rendered. Representing this perimeter requires auxiliary storage and extra clipping computations. This list must then be queried before each new facet's edge is displayed, and the display must be updated if any part of the facet is visible. In contrast, a back-to-front ordering requires no additional storage because the proper occlusion is handled by the drawing order using the painter's algorithm.

---

[1] Along the rows and columns of a discretely sampled image array

[2] This is often called the homogeneous element of the projective coordinate.

5. Reconstruction Issues

The planar-warping equation describes a mapping of the image-space points on a reference viewing plane to image-space points on a desired viewing plane. The underlying assumption is that both the reference and desired images are continuous over their domains. This is not generally the case for typical images. Usually, images are represented by a two-dimensional array of discrete samples. There are many subtle implications of warping sampled images rather than continuous ones. While the warping equation can easily be applied to the discrete coordinates of a sampled image, the likelihood that any sample will map exactly onto a sampling-grid point of the desired image is negligible. In addition to warping to locations off the sampling grid, the points of a reference image will also distribute unevenly over the desired image. The desired image must then be synthesized from this irregular distribution of sampled and reprojected image points.

The process of mapping a sampled image from one image-space to another is called *image resampling*. Conceptually, image resampling constructs a continuous representation of a reference image which is then mapped onto the desired viewing plane using the warping function and resampled to form the final discrete image. When the three-dimensional points represented in the reference image lie on a common plane, as shown in Figure 8, and the scene contents are appropriately band limited, reconstructing a continuous representation is a straightforward application of signal processing theory [Wolberg90]. The same situation occurs when three-dimensional points are constrained to lie along the same rays of different viewing planes, as when reprojecting an arbitrary set of points from the same center-of-projection as shown in Figure 7. The ideal continuous reconstruction of these surfaces is the summation of two-dimensional sinc functions centered at each sample-grid point and scaled by the sample's intensity. Since the sinc function has an infinite extent, local polynomial approximations are often used instead [Mitchell88]. However, when the three-dimensional points visible in an image are not constrained to lie in a plane or share a center-of-projection, the reconstruction of a continuous reference image representation is more involved.

Three-dimensional surfaces can be built up from discrete sub-surfaces, called *surface patches*. The composite of a group of surface patches can represent the reconstruction of a three-dimensional point set. One measure of the continuity of a composite set of surface patches, called *derivative continuity*, is the number of matching terms in the Taylor series expansions at abutting surface patches. When only the first terms of the composite surfaces' Taylor series expansions match along their common boundaries, the surface patches will have the same three-dimensional coordinate values along their abutting regions, and the overall surface has C^0 continuity. When both the first and second terms of the expansion agree, the tangent spaces of the abutting surfaces coincide, and the composite surface will have C^1 continuity.

One approach to reconstructing a continuous function from a sampled reference image is to consider each sample as specifying a surface patch. The basis for these surface patch definitions is typically polynomial. However, this is not a requirement for defining a continuous reconstruction. All that is

necessary is that the basis functions are sufficiently differentiable. For instance, a sinc or a Gaussian basis are both valid representations of surface patches[3].

I will next describe two different methods for constructing a continuous representation of a reference image for use in image warping based on C^0 continuity models that use different surface patch bases. Recent work by Mark [Mark97] has extended these methods to include a model in which C^0 and C^1 continuity alternates throughout the domain. His approach requires a local estimate of the tangent space at each sample point. This requirement can be easily satisfied when the reference images are synthesized using traditional computer graphics techniques, but it is more difficult for acquired reference images. However, there is some promise in this area. It appears that many shape-from-shading [Horn89] and photometric-ratio-based correspondence methods [Wolff94] can be adapted to estimate a surface normal at each image-space point.

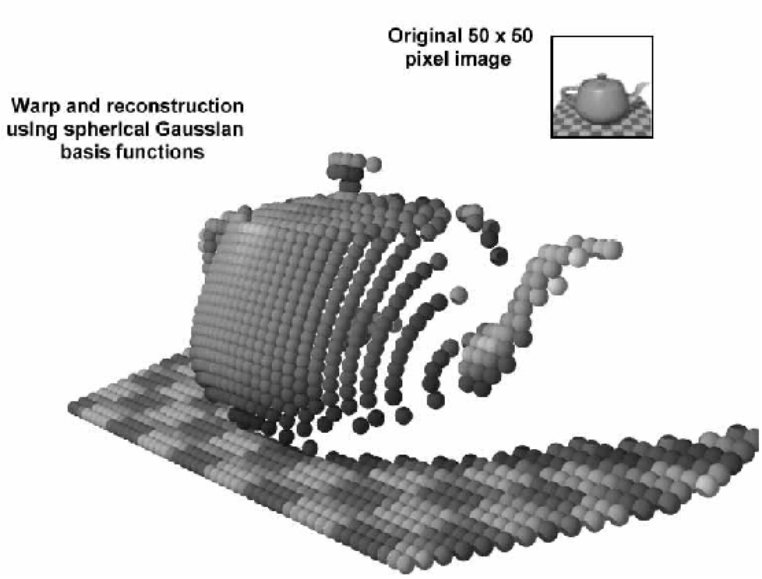

Figure 15: A Gaussian cloud representation of image-space points

The first reconstruction approach uses a spherical Gaussian basis function to represent each sample point. It assumes that every visible point in a reference image represents a three-dimensional spherical cloud density located somewhere along the ray determined by the image point. In three-dimensions this radius could be determined by considering the solid angle represented at each sample point and the distance of the cloud's center from the image's center-of-projection. A two-dimensional equivalent of this radius calculation can, however, be computed directly from the warping equation.

The image-space Gaussian reconstruction method described here is a straightforward adaptation of Heckbert's Elliptical Weighted Average (EWA) filter [Heckbert89], but it is used in a forward-mapping algorithm, similar in spirit to the Splatting algorithm described by Westover [Westover91]. The support of

[3] The use of a sinc or Gaussian basis is somewhat complicated by their infinite extents.

the Gaussian reconstruction function is determined by computing the change in shape of differential circular regions surrounding each sample as they undergo the image warp. One useful measure of the change in shape is provided by the *Jacobian determinant* of the mapping function. The Jacobian determinant is a direct measure of the local change in area induced by a transformation. However, changes in differential area are not necessarily a good indication of the changes in differential shape[4]. The additional assumptions required to address this discrepancy will be discussed shortly. The Jacobian matrix of the planar-warping function given in Equation 12 is

$$\mathbf{J} = \begin{bmatrix} \frac{w_{11}(w_{32}v+w_{33}+w_{34}\delta(u,v))-w_{31}(w_{12}v+w_{13}+w_{14}\delta(u,v))}{(w_{31}u+w_{32}v+w_{33}+w_{34}\delta(u,v))^2} & \frac{w_{12}(w_{31}u+w_{33}+w_{34}\delta(u,v))-w_{32}(w_{11}u+w_{13}+w_{14}\delta(u,v))}{(w_{31}u+w_{32}v+w_{33}+w_{34}\delta(u,v))^2} \\ \frac{w_{21}(w_{32}v+w_{33}+w_{34}\delta(u,v))-w_{31}(w_{22}v+w_{23}+w_{24}\delta(u,v))}{(w_{31}u+w_{32}v+w_{33}+w_{34}\delta(u,v))^2} & \frac{w_{22}(w_{31}u+w_{33}+w_{34}\delta(u,v))-w_{32}(w_{21}u+w_{23}+w_{24}\delta(u,v))}{(w_{31}u+w_{32}v+w_{33}+w_{34}\delta(u,v))^2} \end{bmatrix}$$

Equation 19: Jacobian of the warping equation

The determinant of the Jacobian matrix simplifies to

$$\frac{\partial(u',v')}{\partial(u,v)} = \frac{\det(\mathbf{H}) + \delta(u,v)\det(\mathbf{G})}{t(u,v,\delta(u,v))^3}$$

Equation 20: Jacobian determinant of the warping equation

where

$$\mathbf{H} = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} \text{ and } \mathbf{G} = \begin{bmatrix} w_{11} & w_{12} & w_{14} \\ w_{21} & w_{22} & w_{24} \\ w_{31} & w_{32} & w_{34} \end{bmatrix}$$

Equation 21: Camera-model matrix, H, and the structure matrix, G

The $\mathbf{H}$ component of the Jacobian determinant represents the change in projected area of an infinitesimal region of the reference image due entirely to the changes in the pinhole-camera model since $\mathbf{H} \doteq \mathbf{P}_2^{-1}\mathbf{P}_1$ and $\det(\text{Jacobian}(\mathbf{H}\bar{x})) = \det(\mathbf{H})/t(u,v,0)^3$. The $\mathbf{G}$ component represents the change in projected area due to the three-dimensional structure of the observed image point. This can be seen by letting $\mathbf{H} = \mathbf{I}$, which indicates no change in the pinhole camera model between the reference and desired image. This gives $\frac{\partial(u',v')}{\partial(u,v)} = \frac{1}{(1+\delta(u,v)w_{34})^2}$, indicating that the change in projected area is independent of the point's coordinate on the image plane, but instead it depends entirely on the generalized disparity value at that

---

[4] For instance, consider the mapping $\phi: \{x = 10u, y = 0.1v\}$ with $\frac{\partial(x,y)}{\partial(u,v)} = \text{Determinant}\begin{bmatrix} 10 & 0 \\ 0 & 0.1 \end{bmatrix} = 1$. The Jacobian determinant indicates that the differential area surrounding any sample remains constant; yet the mapping introduces considerable stretching along the u dimension and shrinking in the v dimension.

point. Since the matrices **H** and **G** are constant for a given pair of reference and desired views, they need only be calculated once per warp.

If we make the assumption that the warping function is well represented by a local linear approximation centered at each sample in the reference image with a constant generalized disparity value, then the Jacobian matrix can also be used to estimate the change in shape of a rotationally symmetric reconstruction kernel as follows. An infinitesimal circular region with a radius of dr is described by the expression

$$dr^2 = \begin{bmatrix} du & dv \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix}$$

Equation 22:

When that region undergoes an arbitrary mapping, $\phi: (u, v) \rightarrow (u', v')$, the best linear approximation to the differential mapping is given by the Jacobian matrix

$$\begin{bmatrix} du' \\ dv' \end{bmatrix} = \mathbf{J} \begin{bmatrix} du \\ dv \end{bmatrix}$$

Equation 23:

By substituting Equation 23 into Equation 22, the mapping of the differential circular region can be determined as follows:

$$dr^2 = \begin{bmatrix} du' & dv' \end{bmatrix} (\mathbf{J}^{-1})^T \mathbf{J}^{-1} \begin{bmatrix} du' \\ dv' \end{bmatrix}$$

Equation 24:

which expands to the following conic expression:

$$dr^2 = \frac{(j_{21}^2 + j_{22}^2)du'^2 - 2(j_{11}j_{21} + j_{12}j_{22})du'dv' + (j_{11}^2 + j_{12}^2)dv'^2}{(j_{11}j_{22} - j_{12}j_{21})^2}$$

Equation 25:

This equation describes an ellipse if the Jacobian determinant is positive. Therefore, any circular region centered about a reference image sample will project as an ellipse in the desired image. This property can be used to reconstruct a continuous representation of the reference image prior to resampling. The spatial domain response of any rotationally symmetric filter, such as a Gaussian, can be computed at a sample point in the desired image by evaluating Equation 25 at that point to determine the radius value in the undistorted filter's kernel. The resampling process can be optimized by first computing the extents of the ellipse in the desired image space. In terms of the Jacobian these extents are given by the following expressions:

$$\Delta u' = \pm \sqrt{\frac{dr^2 (j_{11} j_{22} - j_{12} j_{21})^2}{(j_{21}^2 + j_{22}^2) - \frac{(j_{11} j_{21} + j_{21} j_{22})^2}{(j_{11}^2 + j_{12}^2)}}} \qquad \Delta v' = \pm \sqrt{\frac{dr^2 (j_{11} j_{22} - j_{12} j_{21})^2}{(j_{11}^2 + j_{12}^2) - \frac{(j_{11} j_{21} + j_{21} j_{22})^2}{(j_{21}^2 + j_{22}^2)}}}$$

Equation 26:

Another approach to reconstructing a continuous representation of the reference image attempts to fit a bilinear surface patch between any four neighboring grid samples. This representation also has C^0 continuity, but it uses a polynomial basis. First, the individual sample points of the reference image are mapped to the desire image's coordinate system. These warped points will generally not correspond to sample grid positions. The connectivity of the eight-connected neighborhood of each reference sample-point is maintained after the warp. This can be managed by maintaining a buffer of two scan lines while enumerating the reference image in a visibility compatible order. The warped sample points stored in these two buffers can be considered a single strip of patches at the completion of each scan line. A standard polygonal rasterizer can be used to scan convert each patch in the strip. Therefore, this technique can easily take advantage of specialized rasterization hardware if it is available. Once a strip is rasterized, one of the scanline buffers becomes available for storing the mapped values of the next scan line.

Shown below is an example of the same warp using each of the two reconstruction methods discussed.

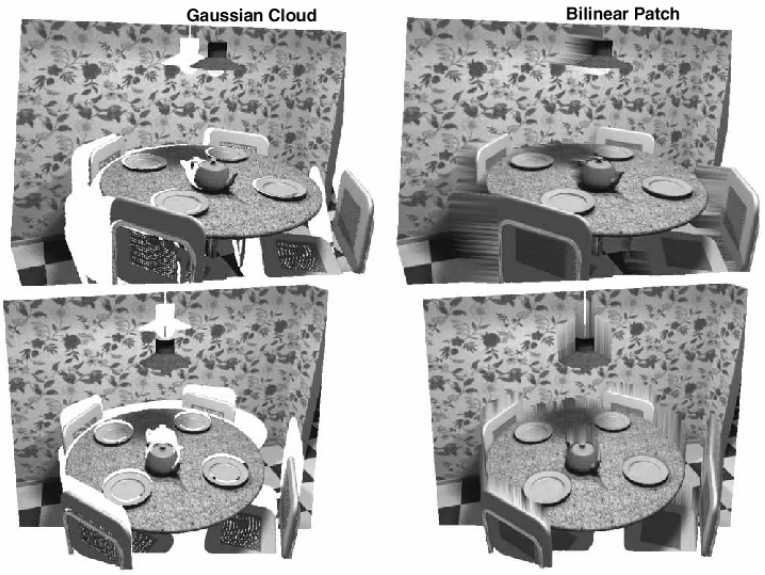

Figure 16: Example image warp using different reconstruction methods

6. Occlusion and Exposure Errors

The image warp can only correctly reproject those scene points visible in the reference image. In some cases, even the visibility of a point does not guarantee its proper reconstruction. These are not

weaknesses of either the image-warping or visibility methods described; they are, instead, inherent limitations of an image-based representation.

The major visual artifacts resulting from these limitations can be classified as one of two cases, *exposure errors* or *occlusion errors*. Exposure errors occur when a background region that should have been occluded is visible in a desired image because of the absence of some foreground element from the reference image. On the other hand, occlusion errors occur in a desired image when an interpolation error in the reconstruction process introduces a false foreground element that covers background regions visible in the actual scene. The choice of reconstruction methods plays a significant role in either amplifying or reducing these errors. However, a reduction in one class of artifact often causes an increase in the other.

Many exposures and occlusions are correct. For instance, when a viewpoint moves toward a foreground object the projection of the object will enlarge in the field-of-view such that it covers adjacent background points. As the viewpoint moves even closer to the foreground object, more of the background is occluded.

Exposure errors and occlusion errors take place either when the underlying assumptions of the reconstruction method are violated, or when there is insufficient information in the image to properly reconstruct the correct surface. These two error sources are closely related. The role of reconstruction kernel is to interpolate the missing gaps between samples. Often the information needed to correctly fill a missing image region is unavailable from the reference image.

Exposure errors are the subtler visual artifact. The region uncovered by a legitimate exposure lends itself to interpretation as a shadow produced by a light source placed at the reference image's center-of-projection. This is particularly noticeable when the exposure occurs along object boundaries. An exposure error occurs when a ray in the desired image passes through this shadow region, allowing some background element to be erroneously seen. Both exposure errors and actual exposures are illustrated below.

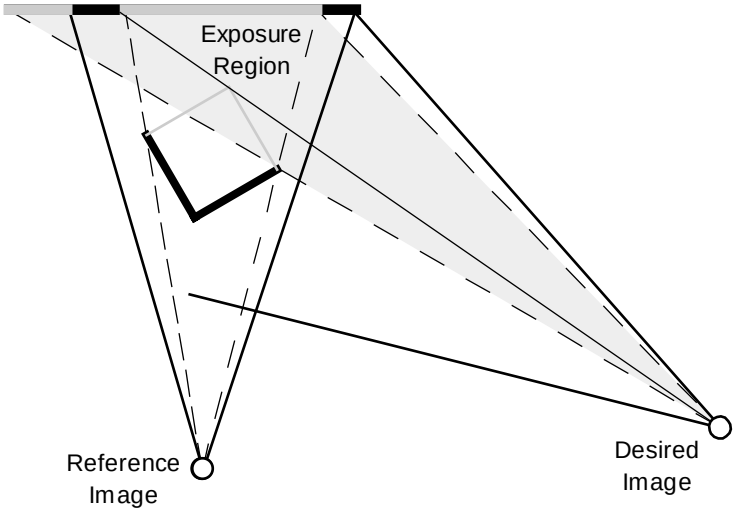

Figure 17: Exposures at occluding boundaries

The actual scene points that are visible from the reference image are shown darkened. The shaded region indicates where an exposure occurs in the desired image. The solid dividing line through the exposure region indicates the boundary between an actual exposure to the right and an exposure error to the left. However, without external information the difference between valid and invalid exposures cannot be resolved.

Exposure errors are most likely to occur at object silhouettes. They occur on smooth surfaces as well as along sharp depth discontinuities. This situation is depicted in Figure 18. Exposure errors occur immediately adjacent to those image points whose ray lies in the observed object's tangent plane. Therefore, as an observer moves to see around an object boundary, she should generally expect to see more of the object rather than any component of the background.

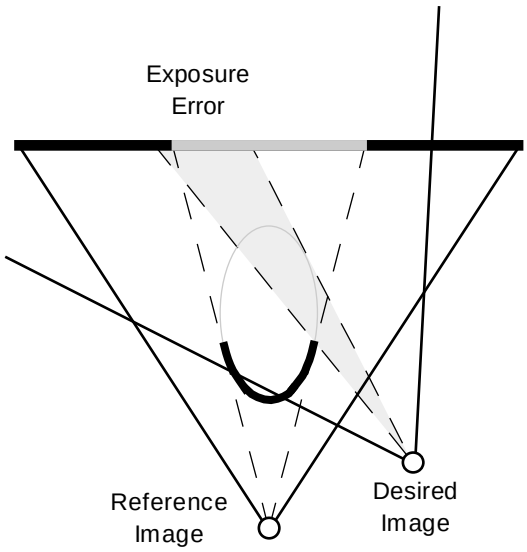

Figure 18: Exposure error on a smooth surface boundary

Merely changing the basis function used in the reconstruction of the reference image can eliminate exposure errors, but it introduces occlusion errors. Consider the example shown in Figure 19 when a polynomial basis, instead of the Gaussian cloud model, is used to approximate the underlying surface.

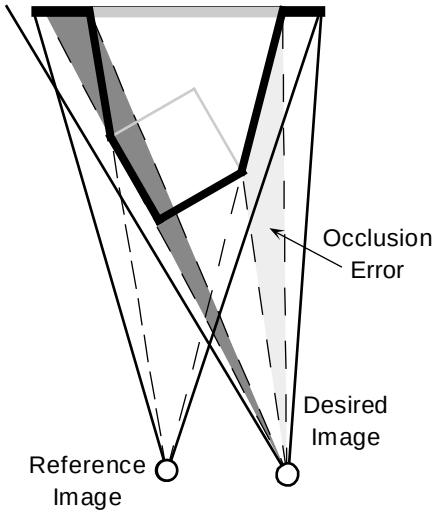

Figure 19: Occlusion errors introduced by polynomial reconstruction

The lighter shaded area indicates the extent of an occlusion error, whereas the darker shaded area represents an actual occlusion. The surface seen along the occlusion error corresponds to an extension of the foreground object's tangent plane. This surface will always enclose the actual surface. However, it is unlikely that any part of this extension will actually represent a point in the environment. The occlusion

error will usually hide valid exposures. Furthermore, since occlusion errors are not adjacent to an actual occlusion, they appear more unnatural than exposure errors.

The continuity of the polynomial interpolation basis causes an excessive estimation of the number of points in the scene. The polynomial interpolation model reconstructs images that are indistinguishable from a model that assumes that all of the points beyond each ray's observed point are occupied. Such a model will not miss any actual scene points, but it will erroneously include scene points that do not exist. The polynomial reconstruction method will, therefore, introduce a disproportionate number of occlusion errors. This can greatly hinder the usefulness of such a model.

In contrast, the Gaussian reconstruction method represents a vacuous estimate of a scene's contents. It assumes that the visible image point is the only scene point located along the ray's extent. Therefore, a Gaussian reconstruction basis will correctly represent all empty regions of an environment, while missing all scene points that are not visible in the reference image. It is, therefore, conservative in its estimate of the occupied volume of space, whereas the polynomial reconstruction makes a conservative estimate of the space's emptiness. Exposure errors should be expected when the Gaussian reconstruction model is used.

The visibility algorithm generally handles valid occlusions. One exception is the case when a scene point from outside the viewing frustum comes into view as a result of the change in the center-of-projection. This situation results in an *invisible occluder error*. A simple example of this case is shown in Figure 20. This problem is a direct result of the limited field-of-view available to a planar-pinhole camera. The use of panoramic pinhole-camera models can remedy this problem.

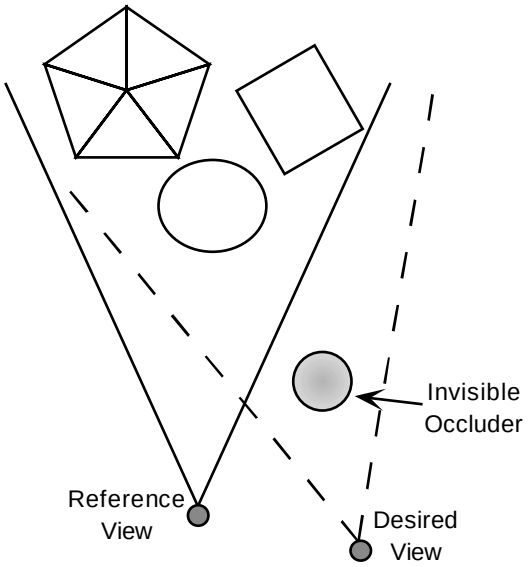

Figure 20: An external exposure error

The two reconstruction methods discussed previously represent two extreme assumptions concerning the structure of space. The use of either of these extreme positions introduces artifacts in the

final rendering. A more accurate reconstruction should combine elements of both methods. However, this might require additional information beyond that which is deducible from the reference image alone. This is an interesting area for future research.

7. Summary

This report has presented a complete example of an image-based method for synthesizing computer graphics. Special mapping functions, called image warps, were derived that enabled arbitrary views of a scene to be generated. This underlying scene was represented by a reference image that was augmented by a scalar value defined at each point, called generalized disparity. An algorithm was also presented to resolve the visibility of the mapped reference points at each coordinate in the desired image. Two methods were presented for reconstructing continuous representations of the warped reference image from these warped points.

8. References

[Anderson82] Anderson, D., *"Hidden Line Elimination in Projected Grid Surfaces,"* **ACM Transactions on Graphics**, October 1982.

[Appel67] Appel, A., *"The notion of quantitative invisibility and the machine rendering of solids,"* **Proceedings of the ACM National Conference**, ACM, New York, October, 1967, pp. 387-393.

[Catmull74] Catmull, E., *"A Subdivision Algorithm for Computer Display of Curved Surfaces,"* Ph.D. Thesis, Department of Computer Science, University of Utah, Tech. Report UTEC-CSc-74-133, December 1974.

[Chen93] Chen, S. E. and L. Williams. *"View Interpolation for Image Synthesis,"* **Computer Graphics** (SIGGRAPH'93 Conference Proceedings), July 1993, pp. 279-288.

[Faugeras92b] Faugeras, O. D., *"What can be seen in three dimensions with an uncalibrated stereo rig?,"* **Proceedings of the 2nd European Conference on Computer Vision**, Springer-Verlag, 1992, pp. 563-578.

[Heckbert89] Heckbert, P. S., *"Fundamentals of Texture Mapping and Image Warping,"* Masters Thesis, Department of EECS, UCB, Technical Report No. UCB/CSD 89/516, June 1989.

[Horn89] Horn, B.K.P., *"Obtaining Shape from Shading Information,"* **Shape From Shading**, Chapter 6, Edited by Berthold K. Horn and Michael J. Brooks, The MIT Press, Cambridge, Massachusetts, 1989.

[Mark97] Mark, W. R., L. McMillan, and G. Bishop**, Proceedings of 1997 Symposium on Interactive 3D Graphics** (Providence, Rhode Island, April 27-30, 1997).

[Mitchell88] Mitchell, D.P. and A.N. Netravali, *"Reconstruction Filters in Computer Graphics,"* **Computer Graphics** (SIGGRAPH'88 Conference Proceedings), Vol. 22, No. 4, August 1988, pp. 221-228.

[Rogers85] Rogers, D.F., **Procedural Elements for Computer Graphics**, McGraw-Hill, New York, NY, 1985.

[Szeliski96] Szeliski, R., "*Video Mosaics for Virtual Environments,*" **IEEE Computer Graphics and Applications**, March 1996, pp. 22-30.

[Westover90] Westover, L. A., "*Footprint Evaluation for Volume Rendering,*" **Computer Graphics** (SIGGRAPH'90 Conference Proceedings), Vol. 24, No. 4, August 1990, pp. 367-376.

[Wolberg90] Wolberg, G., **Digital Image Warping**, IEEE Computer Society Press, Los Alamitos, California, 1990.

[Wolff94] Wolff, L. B., and E. Angelopoulou, *"3-D Stereo Using Photometric Ratios,"* **Proceedings of the European Conference on Computer Vision**, Springer-Verlag, 1994, pp. 247-257.