The Space of All Stereo Images

Steven M. Seitz
Department of Computer Science and Engineering
University of Washington, Seattle, WA
seitz@cs.washington.edu

Abstract

A theory of stereo image formation is presented that enables a complete classification of all possible stereo views, including non-perspective varieties. Towards this end, the notion of epipolar geometry is generalized to apply to multiperspective images. It is shown that any stereo pair must consist of rays lying on one of three varieties of quadric surfaces. A unified representation is developed to model all classes of stereo views, based on the concept of a quadric view. The benefits include a unified treatment of projection and triangulation operations for all stereo views. The framework is applied to derive new types of stereo image representations with unusual and useful properties.

1 Introduction

A stereo pair consists of two images with purely horizontal parallax, that is, every scene point visible in one image projects to a point in the same row of the other. We seek to characterize the space of all stereo images. In particular, suppose that you could construct two sensors that each measure light along an *arbitrary* 2D set of viewing rays and place the resulting measurements in an image. What is the range of light rays and sensor designs that produce a stereo pair?

While the geometric properties of perspective images are well understood, relatively little is known about other types of stereo images. Non-perspective image representations have received burgeoning interest in recent years due to the development of exciting new panoramic sensors [2, 10] and a host of applications in vision and graphics. Furthermore, recent results [7] have demonstrated the existence of *stereo panoramas* [12] that enable an observer to achieve a 360° depth perception. Processing such images with stereo algorithms enables a panoramic scene reconstruction from a single image pair [7, 17], a key capability that is not possible with perspective images.

Inspired by this previous work in multiperspective imaging, we seek to identify *all* of the ways in which stereo images may be formed. Towards this end, this paper makes the following contributions:

- Epipolar geometry is extended to multiperspective images. The main result is that three varieties of epipolar geometry exist, corresponding to families of planes, hyperboloids, and hyperbolic-paraboloids.
- It is shown that all stereo images represent rays lying on quadric surfaces. Based on this analysis, a complete classification of stereo images is derived.
- A unified representation is introduced to model all classes of stereo images, based on the concept of a quadric view. The benefits include a unified treatment of projection and triangulation operations for all stereo image varieties.
- A recipe book for generating stereo images is provided. In order to demonstrate the power of this framework, we show how all previous stereo images can be systematically constructed, and present three new varieties with interesting properties. Real examples of one of these representations, which we call *stereo cyclographs*, are shown.
- It is shown that the only three-view epipolar geometry that exists is the planar variety.

These results are perhaps counter-intuitive in two respects. On one hand, they demonstrate that we can potentially fuse images that have multiple centers of projection. This is surprising, in that our visual system is clearly designed for processing two single-perspective images. On the other hand, it is surprising that so *few* varieties of stereo views exist. Out of all possible 2D subsets of the 5D set of rays, only three varieties satisfy the stereo constraint!

Concurrent with this research, Pajdla [11] independently obtained a similar classification of epipolar surfaces (the first bullet in the above list), but specialized for the case of cameras that capture lines rather than rays.

The remainder of the paper is structured as follows. Section 2 introduces a generalized model of image formation and defines the stereo constraint geometrically. Section 3 presents our classification of stereo images and their associated epipolar geometries. Section 4 introduces the concept of a quadric view and derives a unified mathematics

0-7695-1143-0/01 $10.00 © 2001 IEEE

for stereo imaging. The stereo cyclograph is introduced in Section 5 and example images are shown.

2 Stereo Imaging

2.1 A Generalized Imaging Model

Light in any environment flows along straight lines in space[1]. From the viewer's perspective, it is convenient to model light flow in terms of the distribution of light that is received at each point and from each direction in the environment, i.e., along each *light ray* $R = (x, y, z, \theta, \phi)$. The set of all light rays defines a five-dimensional set that we call *ray space* and denote $\mathcal{P}$. The light energy flowing along these rays can be represented as a function $P : \mathcal{P} \rightarrow \mathcal{E}$, where $\mathcal{E}$ represents light energy[2]. P is known as the *plenoptic function* [1].

We use the term *view* to denote any two-dimensional subset of light rays:

> A *view* is any function $V : \mathcal{D} \rightarrow \mathcal{P}$ on a surface $\mathcal{D}$ that maps image coordinates (u, v) to light rays.
>
> An *image* I measures the light observed for a view, i.e., $I = P \circ V$.

Conceptually, a view encodes imaging *geometry*—the distribution of light rays measured by a camera, whereas an image represents *photometry*—light energy radiated from the environment toward the camera. In this paper, it is implicit that every image I has a corresponding view V. Accordingly, any terms that we define for views, e.g., field of view, stereo pair, etc., will also apply to images.

These definitions are very broad, emcompassing most projection models used in the computer vision literature, including single point perspective, orthographic, weak perspective, and affine models [8]. In addition, non-perspective mosaic representations fit this definition, including single-perspective panoramas [2], multiperspective panoramas [19], manifold mosaics [13], multiple-center-of-projection images [14], stereo panoramas [12], concentric mosaics [16], omnivergent images [17], and 360 × 360 panoramas [10][3].

The following terminology and notation will be useful. The *field of view* of a view V, denoted $FOV(V)$ is defined to be the range of V, i.e., $FOV(V) = \{V(u, v) \mid (u, v) \in \mathcal{D}\}$. We say a point $\mathbf{X}$ is *within the field of view* of V, denoted $\mathbf{X} \in FOV(V)$, if $\mathbf{X}$ lies along some ray in $FOV(V)$.

[1] For the purposes of this paper, relativistic effects due to the curvature of space are ignored.

[2] In RGB color space, $\mathcal{E}$ is isomorphic to $\Re^3$. More generally, $\mathcal{E}$ can be a function space representing the light spectrum.

[3] Note that this definition of an image requires that every image point corresponds to a unique light ray–point spread functions, for instance, are not modeled.

We write $\mathbf{X} \in R$ when a 3D point $\mathbf{X}$ lies along a ray R. The intersection of two rays R_1 and R_2 is written $R_1 \cap R_2$. In general, this intersection can be a point, a ray, or be empty.

The *projection* of a point $\mathbf{X}$ to a view V is defined to be the set of tuples $(u, v) \in \mathcal{D}$ such that $\mathbf{X} \in V(u, v)$. The projection of a set of points $\mathcal{X}$ is defined to be the union of the projections of points in $\mathcal{X}$.

We use the notation $\mathcal{D}^u$ to denote column u in a view, and $\mathcal{D}_v$ to denote row v. We say a view V is *u-continuous* if, for every value of v, $V(u, v)$ is a continuous function of u and connected sets of points in $FOV(V)$ project to connected subsets of $\mathcal{D}_v$. *v-continuity* is defined analogously.

2.2 Multiperspective Imaging

Pin-hole camera models have the property that all rays pass through a single point in 3D space, known as the *center of projection*. Driven by applications in graphics [19, 14, 16] and stereo matching [7, 17], several researchers have proposed view representations that do not have a single distinguished center of projection. Rather, these *multiperspective* views capture rays that emanate from different points in space. We define the *generator* of a view to be the set of "camera centers" for that view:

> The *generator* of a view V is defined to be $\{(x\ y\ z) \mid (x, y, z, \theta, \phi) \in FOV(V)\}$.

2.3 The Stereo Constraint

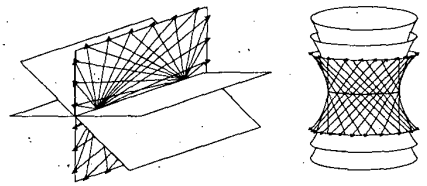
A stereo pair consists of two u-continuous views with purely horizontal parallax, that is, every scene point visible in one view projects to a point in the same row of the other view. Images satisfying this property can be fused by human observers to produce a depth effect and are amenable to processing by computational stereo algorithms.

More formally, we say that two u-continuous views V_1 and V_2 satisfy the *stereo constraint* if the following property holds:

> The rays $V_1(u_1, v_1)$ and $V_2(u_2, v_2)$ intersect only if $v_1 = v_2$.

Any two such views are referred to as *stereo views* or a *stereo pair*.

It is often the case that views in a stereo pair will overlap only partially, i.e., there are points in the scene that lie within the field of view of one view but not the other. Such points cannot be "fused" and we therefore limit our analysis to regions of the scene within the field of view of both views. We also ignore points that are imaged with the same ray in both views, since such points do not provide stereo depth cues. For instance, two identical images from the same perspective camera viewpoint satisfy our definition of



(a) (b)

Figure 1: Epipolar Geometry. (a) Rays in two perspective views lie on a pencil of epipolar planes. (b) Rays in stereo panoramas and 360 × 360 panoramas lie on epipolar hyperboloids.

a stereo pair, but yet do not provide parallax cues that can be used to infer depth. Accordingly, we define the stereo viewing region as follows

> A point $\mathbf{X}$ in 3D space is *stereo viewable* from V_1 and V_2 if there exist rays $R_1 \in FOV(V_1)$ and $R_2 \in FOV(V_2)$ such that $R_1 \cap R_2 = \mathbf{X}$.

The set of all points $\mathbf{X}$ that are stereo viewable from two views V_1 and V_2 is referred to as the *stereo viewable space*, denoted $\mathcal{V}_1^2$. Note that there are generally points in the stereo viewable space that are occluded in one or both images. Occlusions will not affect our analysis, however, since the stereo constraint is defined by the distribution of rays, independent of visibility.

2.4 Central Perspective Stereo

The classical example of a stereo pair consists of two planar perspective views where the second view has been translated horizontally from the first[4]. The fact that two such images have horizontal parallax is important both for human perception and computational stereo vision.

It has long been known that a perspective stereo pair (V_1, V_2) obeys a very special geometric constraint: the rays in both views all lie on a pencil of planes, as shown in Fig. 1(a). In particular, the rays in row v of V_1 sweep out a surface that lies on a plane in the scene, denoted P_v. Similarly, row v of V_2 sweeps out a second surface that also lies on P_v. P_v is known as the *epipolar plane* of row v of the image.

Note that *any* pair of perspective images defines a set of epipolar planes, i.e., epipolar geometry apply to all perspective images, not just to stereo pairs. In fact, it has been shown that epipolar geometry applies to panoramic views that are defined on cylindrical [9] or spherical image domains [18]. In general, it can be seen that any pair of *central views*, i.e., views that are each generated by a single point, can be partitioned into rays that lie on epipolar planes–as shown in Fig. 1(a), this condition is independent of the shape of the imaging surface and depends only on the positions of the generators (camera centers).

Importantly, **any pair of views that define epipolar planes can be converted into a stereo pair by reparameterization of the domain $\mathcal{D}$.** This procedure is known as *rectification* in the stereo literature, and can be applied simply by creating a new domain $\mathcal{D}'$ of the same topology as $\mathcal{D}$ such that rays in the same epipolar plane have the same v coordinate in $\mathcal{D}'$.

2.5 Multiperspective Stereo

While the epipolar condition (referred to as *epipolar geometry*) is well-understood for perspective views, it is not clear what it means for two *multiperspective* views to have epipolar geometry. In particular, can stereo pairs be created from views that do not have a single center of projection? Recent work has shown that multiperspective stereo pairs do indeed exist. For instance, Ishiguro et al. [7] and Peleg et al. [12] described how *stereo panoramas* could be produced by capturing a set of viewing rays passing through a circle. Interestingly, stereo panoramas can be shown to have purely horizontal parallax [17], but, unlike perspective views, allow for a 360° field of view.

Do stereo panoramas have epipolar geometry? Clearly, the rays from stereo panoramas do not all lie on a pencil of epipolar planes. Note that each scanline of the left view is formed by rotating a single ray around a circle, sweeping out a hyperboloid $\mathcal{S}_1$ as shown in Fig. 1(b). The corresponding scanline of the right view defines a second hyperboloid $\mathcal{S}_2$. Note that $\mathcal{S}_1 = \mathcal{S}_2$, ensuring that the left and right scanlines correspond to the same subset of the scene. In the same way that perspective rays lie on epipolar planes, rays from stereo panoramas can be shown to lie on "epipolar hyperboloids."

3 A Classification of Stereo Views

What is the space of all stereo views? In this section we derive a full categorization of pairs of views that satisfy the stereo constraint.

3.1 Epipolar Surfaces

Let V_1 and V_2 be two stereo views, and consider a specific row v. Since both views are u-continuous, the light rays $V_i(u, v)$ sweep out a *ruled surface* $\mathcal{S}_i$. We are interested in the subset of $\mathcal{S}_1$ and $\mathcal{S}_2$ that is stereo viewable. In particular, define the *epipolar surface* for row v to be $\mathcal{S}_1^2(v) = \mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathcal{V}_1^2$. In the case of perspective views, for example, $\mathcal{S}_1^2(v)$ is planar for all v.

In the general case, observe that $\mathcal{S}_1^2(v)$ contains two families of straight lines–it lies on a *doubly-ruled surface* (see Fig. 1). We may therefore characterize the space of stereo

[4]More specifically, "horizontal" means in the direction of a row on the image plane of the first view.

generator	double-ruled quadric
point	plane
line	plane, hyperboloid, hyperbolic paraboloid
ellipse	hyperboloid
parabola	hyperbolic paraboloid
hyperbola	hyperboloid, hyperbolic paraboloid

Table 1: Any quadric view may be generated by moving a sensor along a conic path. This table shows the doubly-ruled quadrics for each conic generating path.

views by classifying the set of all doubly-ruled epipolar surfaces. Fortunately, the latter set can be explicitly characterized. In particular, the only doubly-ruled surfaces are the plane, hyperboloid, and hyperbolic paraboloid [6]. While this classical result applies to algebraic surfaces ruled by lines, with some care it can be extended to the case of non-parametric surfaces ruled by rays or line segments. Due to space limitations, we present the proof of this result in a companion technical report [15].

We therefore have the following result:

> **Stereo Classification Theorem**: V_1 and V_2 are stereo views only if for every row v, $\mathcal{S}_1^2(v)$ lies on a plane, hyperboloid, or hyperbolic paraboloid.

This result demonstrates that we can potentially fuse images that have multiple centers of projection, a result that may seem counter-intuitive. Also surprising is that so *few* varieties of stereo views exist–out of all possible 2D subsets of the 5D set of rays, only three varieties satisfy the stereo constraint.

Note that this theorem concerns the distribution of rays in space, but does not specify the exact parameterization. Different (u,v) parameterizations of the same set of rays will produce distinct images. We could further generalize the space of stereo images by allowing any parameterization of rays so that epipolar surfaces map to arbitrary curves (or 1D point sets) rather than horizontal lines. These generalized stereo images could always be *rectified* to produce images with horizontal epipolar lines, much in the way that perspective image pairs are rectified. Regardless of parameterization, however, the plane, hyperboloid, and hyperbolic paraboloid are the only types of epipolar surfaces.

We use the term *quadric view* to denote a view having the property that all rays in each row lie on a quadric surface.

3.2 Three- and N-View Stereo

The stereo constraint is easily generalized to three or more views by requiring that the two-view stereo constraint be satisfied for every pair of views. A direct consequence of this condition is that all three views must share the same set of epipolar surfaces. In particular, each row must sweep out a *triply-ruled surface*. Since every triply-ruled surface is also doubly-ruled, the epipolar surfaces must lie on planes, hyperboloids, or parabolic paraboloids, by the Stereo Classification Theorem. Of these only the plane is a triply-ruled surface (it has an infinite number of rulings). It follows that the only epipolar geometry for three or more views is the planar variety. The *pushbroom panorama* described in Section 3.3 is an example of an image representation that can be used to form stereo triplets, quadruplets, etc.

image representation	generator	epipolar surfaces
perspective	point	pencil of planes
stereo panorama [7, 12, 17]	circle	half-hyperboloids
360 × 360 [17, 10]	circle	hyperboloids
spherical omnivergent [17]	sphere	pencil of planes
pushbroom panorama	line	pencil of planes
stereo cyclograph	ellipse	half-hyperboloids
parabolic panorama	parabola	hyper. paraboloids

Table 2: Classification of known stereo view varieties. The last three variants are introduced in this paper.

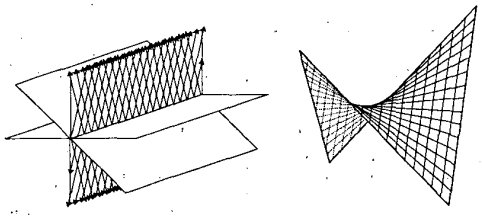
3.3 Generating Stereo Views

We now turn to the problem of how stereo views may be captured. Whereas perspective views can be imaged with a small CCD array, the same is not necessarily true for multiperspective views. In general, multiperspective views require placing sensors on a generating path or surface. In principle, this generator could be arbitrarily complex. Due to their restrictive form, however, stereo views are generated with relative ease.

In particular, any quadric view may be generated by capturing rays from a conic path. For instance, a hyperboloid is generated by a line, ellipse, or hyperbola. Conversely, any conic is the generator for *some* quadric surface. Table 1 summarizes the doubly-ruled quadrics corresponding to each conic generator. This table may be used as a recipe book for generating stereo pairs. For instance, suppose you wish to generate a stereo pair with hyperbolic paraboloid epipolar geomery. Table 1 indicates that you should capture the appropriate rays by moving a camera or placing sensors on a line, parabola, or hyperbola.

Because our classification is comprehensive, we can characterize all previously proposed varieties of stereo images in terms of their generators and corresponding families of quadric epipolar surfaces (Table 2). In addition, the classification suggests new varieties of interest, three of which are listed in the table. The *pushbroom panorama* and *parabolic image* are described briefly in this section, and the *stereo cyclograph* is presented in detail in Section 5.

A pushbroom panorama is generated by translating a camera to create a sequence of images with purely horizontal parallax. Suppose we translate a perspective camera parallel to the image-plane u-axis, and extract column i of pixels from every input image. Placing these columns into successive columns of a new image I_i creates a new image



(a) (b)

Figure 2: Geometry of a pushbroom panorama (a) and a parabolic image (b).

representation which we call a *pushbroom panorama*. Selecting different subsets of rays from each view produces other varieties of pushbroom panoramas. For instance, capturing a cone of rays symmetric about the translation direction generates the omnidirectional pushbroom panorama shown in Fig. 2(a). These image representations are generalizations of pushbroom cameras [5]. Any two pushbroom panoramas I_i and I_j form a stereo pair with rays lying on a pencil of epipolar planes, as shown in Fig. 2(a). An example application of pushbroom panoramas is to create a stereo panorama of an entire city street or landscape.

The second new variety of stereo pair proposed in this paper is generated by moving a camera on a parabolic path. As indicated in Table 1, this strategy can be used to create views whose rows lie on hyperbolic paraboloids, as shown in Fig. 2(b). Two such images that share the same set of hyperbolic paraboloids form a stereo pair. While of theoretical interest, parabolic stereo pairs have a rather limited stereo viewable region with the wedge shape shown in Fig. 2(b). Consequently, a stereo effect is achieved only for objects near the generator, making this image representation less attractive compared with the alternatives in Table 2.

4 Quadric Imaging Representation

Equipped with a classification of stereo pairs, we now turn to the question of how to represent and operate on such views. The objective is to obtain a unified approach that applies to every type of stereo pair, including multiperspective views. Since all stereo pairs are made up of rays lying on quadric surfaces, it is convenient to cast our formulation in the framework of projective geometry.

4.1 Terminology

We denote points in the scene as 4D homogeneous column vectors $\mathbf{X} = [x\ y\ z\ 1]^T$ and planes as 4D homogeneous row vectors $\mathbf{P} = [a\ b\ c\ d]$. Point $\mathbf{X}$ is on plane $\mathbf{P}$ if and only if

$$\mathbf{PX} = 0 \tag{1}$$

Any quadric surface $\mathcal{Q}$ is represented as the set of solutions to a quadratic equation:

$$\mathcal{Q} = \{\mathbf{X} \mid \mathbf{X}^T\mathbf{QX} = 0\} \tag{2}$$

where $\mathbf{Q}$ is a symmetric 4×4 matrix.

4.2 Quadric Representation

Let V be a u-continuous view. Each row of V defines a ruled surface S swept out by the rays in that row. If every row of V sweeps out a quadric surface, V is a quadric view. Suppose further that V is also v-continuous with the property that every column also defines a quadric surface. We use the term *bi-quadric view* to denote such a view V. We will devote special attention to bi-quadric views since most of the image representations of interest fit into this category.

Any bi-quadric view is defined by a two parameter family of quadric surfaces. Let $\mathcal{Q}_v$ denote the quadric surface corresponding to row v and $\mathcal{Q}_u$ the quadric surface corresponding to column u of V. The ray $V(u, v)$ is represented as the space of solutions to the following equations:

$$\mathbf{X}^T\mathbf{Q}_v\mathbf{X} = 0 \tag{3}$$
$$\mathbf{X}^T\mathbf{Q}_u\mathbf{X} = 0 \tag{4}$$

plus any number of *field of view* constraints of the form:

$$\mathbf{X}^T\mathbf{Q}_i\mathbf{X} \geq 0 \tag{5}$$

Here $V(u, v)$ is defined *implicitly* to be the set of all $\mathbf{X}$ that satisfy Eqs. (3-5). For the special case that $\mathcal{Q}_u$ is a plane for every value of u, Eq. (4) may be more conveniently written:

$$\mathbf{P}_u\mathbf{X} = 0 \tag{6}$$

and similarly for Eqs. (3) and (5).

For example, consider the case of a planar perspective view V. The perspective projection equations may be written: $\mathbf{u} = \mathbf{\Pi X}$, where $\mathbf{\Pi}$ is the 3×4 projection matrix for that view, and $\mathbf{u} = [su\ sv\ s]^T$. Using the notation that $\mathbf{\Pi}_i$ is the ith row of $\mathbf{\Pi}$, we can rewrite these equations as follows

$$(\mathbf{\Pi}_2 - v\mathbf{\Pi}_3)\mathbf{X} = 0 \tag{7}$$
$$(\mathbf{\Pi}_1 - u\mathbf{\Pi}_3)\mathbf{X} = 0 \tag{8}$$

plus field of view constraints

$$(\mathbf{\Pi}_2 - v_{min}\mathbf{\Pi}_3)\mathbf{X} \geq 0 \tag{9}$$
$$-(\mathbf{\Pi}_2 - v_{max}\mathbf{\Pi}_3)\mathbf{X} \geq 0 \tag{10}$$
$$(\mathbf{\Pi}_1 - u_{min}\mathbf{\Pi}_3)\mathbf{X} \geq 0 \tag{11}$$
$$-(\mathbf{\Pi}_1 - u_{max}\mathbf{\Pi}_3)\mathbf{X} \geq 0 \tag{12}$$

Eqs. (7-12) represent the ray $V(u, v)$ as the intersection of two planes and the viewing frustum, as shown in Fig. 3. $\mathbf{\Pi}_2 - v\mathbf{\Pi}_3$ is the vector corresponding to the plane passing through the camera center and row v of the image plane. Similarly, $\mathbf{\Pi}_1 - u\mathbf{\Pi}_3$ represents the plane from the camera center through colum u of the image plane.

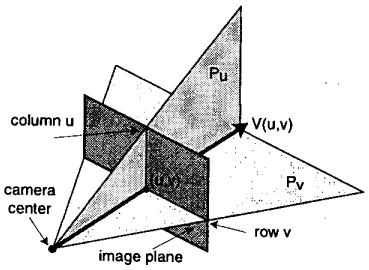


Figure 3: Bi-quadric Image Representation. A perspective view is represented by the two families of planes (quadrics, more generally) $\mathbf{P}_u$ and $\mathbf{P}_v$ and field of view constraints. Each ray in the field of view is specified by the intersection of two of these planes.

As another example, consider the case of a 360 × 360 panorama [10]. As originally described in [17], this variety of stereo view is generated by moving a camera along a circle and capturing a planar strip of rays that is normal to the circle radius. The result is a panoramic image with a 360° horizontal and vertical field of view. Since opposite edges of the image are identified, the domain of the view is a torus.

A 360 × 360 panorama can be represented in our framework by a family of hyperboloids and planes as follows

$$\mathbf{P}_u = [\cos(u) \quad \sin(u) \quad 0 \quad -1] \tag{13}$$

$$\mathbf{Q}_v = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -v & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{14}$$

$$\begin{bmatrix} \sin(u) & \cos(u) & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{X} \geq 0 \tag{15}$$

The field of view (inequality) constraints specify which set of rays are captured along the generating circle for a given image. In particular, this view captures only "forward" rays in the direction of motion around the circle, from the top half of the hyperboloid. The corresponding image generates the top half of a 360 × 360 panorama. The bottom half may be generated with a similar set of constraints.

4.3 Quadric Projection

The camera projection operator maps 3D points $\mathbf{X}$ into view coordinates (u, v). In the case of bi-quadric views, the projection operation maps 3D points onto two families of quadric surfaces, indexed by u and v. Given $\mathbf{X}$, u and v are computed by first checking that $\mathbf{X}$ satisfies the field of view constraints, and then solving Eqs. (4) and (3) for $\mathbf{Q}_u$ and $\mathbf{Q}_v$ respectively. Note that these equations are linear in the elements of $\mathbf{Q}_u$ and $\mathbf{Q}_v$. If $\mathbf{Q}_u$ and $\mathbf{Q}_v$ depend linearly on u and v then the solution of u and v are also linear.

In the case of a 360 × 360 panorama, for example, substituting Eqs. (14) and (13) into Eqs. (3) and (6) yields

$$v = \frac{x^2+y^2-1}{z^2} \qquad u = \pm \arccos\left(\frac{x \pm y\sqrt{x^2+y^2-1}}{x^2+y^2}\right)$$

A unique value of u may be obtained by enforcing the field of view constraints (Eqs. (15)).

4.4 Quadric Triangulation

Let (u_1, v) in V_1 and (u_2, v) in V_2 be projections of the same scene point $\mathbf{X}$. The problem of computing $\mathbf{X}$ from its projections in two images is known as *triangulation*. The triangulation of (u_1, v) and (u_2, v) is computed by solving the following equations for $\mathbf{X}$

$$\mathbf{X}^T\mathbf{Q}_v\mathbf{X} = 0 \tag{16}$$

$$\mathbf{X}^T\mathbf{Q}_{u_1}\mathbf{X} = 0 \tag{17}$$

$$\mathbf{X}^T\mathbf{Q}_{u_2}\mathbf{X} = 0 \tag{18}$$

Once again, Eqs. (17) and (18) may be replaced with $\mathbf{P}_{u_1}\mathbf{X} = 0$ and $\mathbf{P}_{u_2}\mathbf{X} = 0$, respectively, when $\mathcal{Q}_{u_1}$ and $\mathcal{Q}_{u_2}$ are planar.

For the case of a 360×360 panorama, the triangulation of image points (u_1, v) and (u_2, v) is obtained by substituting Eqs. (14) and (13) into Eqs. (16-18). After applying the field of view constraints (Eqs. (15)) the following expression for $\mathbf{X}$ is obtained:

$$\mathbf{x} = \left[\frac{\sin(u_2)-\sin(u_1)}{\sin(u_2-u_1)} \quad \frac{\cos(u_1)-\cos(u_2)}{\sin(u_2-u_1)} \quad \frac{1}{\sqrt{v}}\tan\left(\frac{u_1-u_2}{2}\right) \quad 1\right]^T$$

4.5 The Quadric Fundamental Matrix

Let V_1 and V_2 be a stereo pair. By definition, for two points (u_1, v_1) in V_1 and (u_2, v_2) in V_2 to be in correspondence, v_1 and v_2 must be identical. This constraint is easily encoded in terms of the fundamental matrix equation which has a very simple form for all stereo views:

$$[u_1 \; v_1 \; 1]\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = 0 \tag{19}$$

5 Stereo Cyclographs

An interesting type of panorama may be created by moving an *inward*-facing camera 360° on an ellipse around an object of interest, capturing a column of pixels from each image, and stacking these columns side by side into a new panorama. This construction resembles an approach used by archaeologists to create unwrapped "cyclograph"[5] images of ancient pottery [4]. Using the concepts from Section 3, it can be seen that this construction generates a bi-quadric image, where the rows of the image sweep out hyperboloids and the columns define planes.

In order to create stereo cyclograph images, we placed a video camera on a rotating platform and captured a sequence of images while the camera rotated in a circle around the object. The camera was pointed inward towards the subject (a person's head). The entire image sequence

[5] Cyclographs are also referred to as *periphotographs* or *rollout photographs*.

contains just over 24 seconds of video. If the frames are stacked one on top of the other, they define an x-y-θ volume, similar to the x-y-t volumes introduced by Bolles et al. [3]. Cyclograph images correspond to y-θ slices through the volume, as shown in Fig. 4(a). In particular, a stereo cyclograph is formed from any two y-θ slices that are symmetric about the slice $x = x_0$ of rays that pass through the axis of rotation.

Fig. 5 shows three stereo cyclograph pairs constructed in this manner. Since all rays in a cyclograph are tangent to a cylinder, the interior of the cylinder is not visible. Hence, only the subject's nose, glasses, and pony-tail are visible in the top row of Fig. 5(b) and the rest of the object is seamlessly removed. As the radius of the cylinder is decreased, more of the subject's face comes into the field of view (middle row), until the head is fully visible (bottom row), creating an interesting effect. The bottom pair of images can be fused stereoscopically.

Stereo cyclographs enable viewing "all sides" of an object at once, in stereo. In addition to visualization, cyclographs have potential benefits for shape reconstruction. Existing stereo techniques can be applied to the bottom pair of images in Fig. 5 to obtain a nearly completed 3D head model from a single image pair. The entire set of cyclographs could be used for structure-from-motion or multibaseline stereo processing, perhaps using new algorithms that exploit the unique structure of cyclograph sequences.

6 Conclusion

A theory of stereo image formation was presented that provides a complete classification of all possible stereo views. A main result was that only three varieties of epipolar surfaces exist, corresponding to rays lying on doubly-ruled quadrics. A second important contribution was a novel representation called a *quadric view* that provides a unified mathematical treatment of all stereo views. This framework was used to derive projection and triangulation operations that apply equally to both perspective and multiperspective image representations. The framework was shown to subsume all previous stereo image representations and also enabled deriving three new types of stereo image representations with unusual and useful properties. One of these representations, called the *stereo cyclograph*, was experimentally demonstrated and offers compelling advantages for visualization and 3D reconstruction tasks.

Acknowledgements

We would like to thank Microsoft Research for providing access to their Cyberscan platform, and Sashi Raghupathy, Zicheng Liu, and Eric Hanson for their assistance in acquiring images. The support of Intel Corporation, Microsoft Corporation, and the National Science Foundation under grant IIS-9984672 is gratefully acknowledged.

References

[1] E. H. Adelson and J. R. Bergen. *The Plenoptic Function and the Elements of Early Vision*. MIT Press, Cambridge, MA, 1991.

[2] S. Baker and S. Nayar. A theory of single-viewpoint catadioptric image formation. *Int. J. of Computer Vision*, 35(2):175–196, 1999.

[3] R. C. Bolles, H. H. Baker, and D. H. Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *Int. J. of Computer Vision*, 1(1):7–55, 1987.

[4] A. Davidhazy. Principles of peripheral photography. *Industrial Photography*, Jan. 1987. (http://www.rit.edu/andpph/text-peripheral-basics.html).

[5] R. Gupta and R. I. Hartley. Linear pushbroom cameras. *IEEE Trans. Pattern Analysis and Machine Intell.*, 19(9):963–975, 1997.

[6] D. Hilbert and S. Cohn-Vossen. *Geometry and the Imagination*. AMS Chelsea, Providence, RI, 1991.

[7] H. Ishiguro, M. Yamamoto, and S. Tsuji. Omni-directional stereo. *PAMI*, 14(2):257–262, February 1992.

[8] J. J. Koenderink and A. J. van Doorn. Affine structure from motion. *J. Opt. Soc. Am. A*, 8:377–385, 1991.

[9] L. McMillan and G. Bishop. Plenoptic modeling. In *Proc. SIGGRAPH 95*, pages 39–46, 1995.

[10] S. K. Nayar and A. Karmarkar. 360 x 360 mosaics. In *Proc. Computer Vision and Pattern Recognition Conf.*, pages 388–395, 2000.

[11] T. Pajdla. Epipolar geometry of some non-classical cameras. In *Proceedings of Computer Vision Winter Workshop*, pages 223–233. Slovenian Pattern Recognition Society, 2001.

[12] S. Peleg and M. Ben-Ezra. Stereo panorama with a single camera. In *Proc. Computer Vision and Pattern Recognition Conf.*, pages 395–401, 1999.

[13] S. Peleg and J. Herman. Panoramic mosaics by manifold projection. In *Proc. Computer Vision and Pattern Recognition Conf.*, pages 338–343, 1997.

[14] P. Rademacher and G. Bishop. Multiple-center-of-projection images. *Proceedings of SIGGRAPH 98*, pages 199–206, July 1998.

[15] S. M. Seitz. The space of all stereo images. Technical report UW-CSE-01-04-01, University of Washington, Seattle, Washington, April 2001.

[16] H.-Y. Shum and L.-W. He. Rendering with concentric mosaics. *Proceedings of SIGGRAPH 99*, pages 299–306, 1999.

[17] H.-Y. Shum, A. Kalai, and S. M. Seitz. Omnivergent stereo. In *Proc. 7th Int. Conf. on Computer Vision*, pages 22–29, 1999.

[18] T. Svoboda, T. Pajdla, and V. Hlavac. Epipolar geometry for panoramic cameras. In *Proc. 5th Eur. Conf. on Computer Vision*, pages 218–232, 1998.

[19] D. N. Wood, A. Finkelstein, J. F. Hughes, C. E. Thayer, and D. H. Salesin. Multiperspective panoramas for cel animation. In *Proc. SIGGRAPH 97*, pages 243–250, 1997.

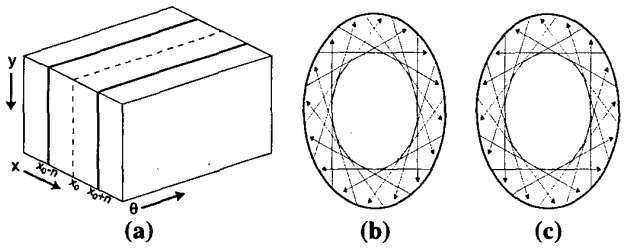
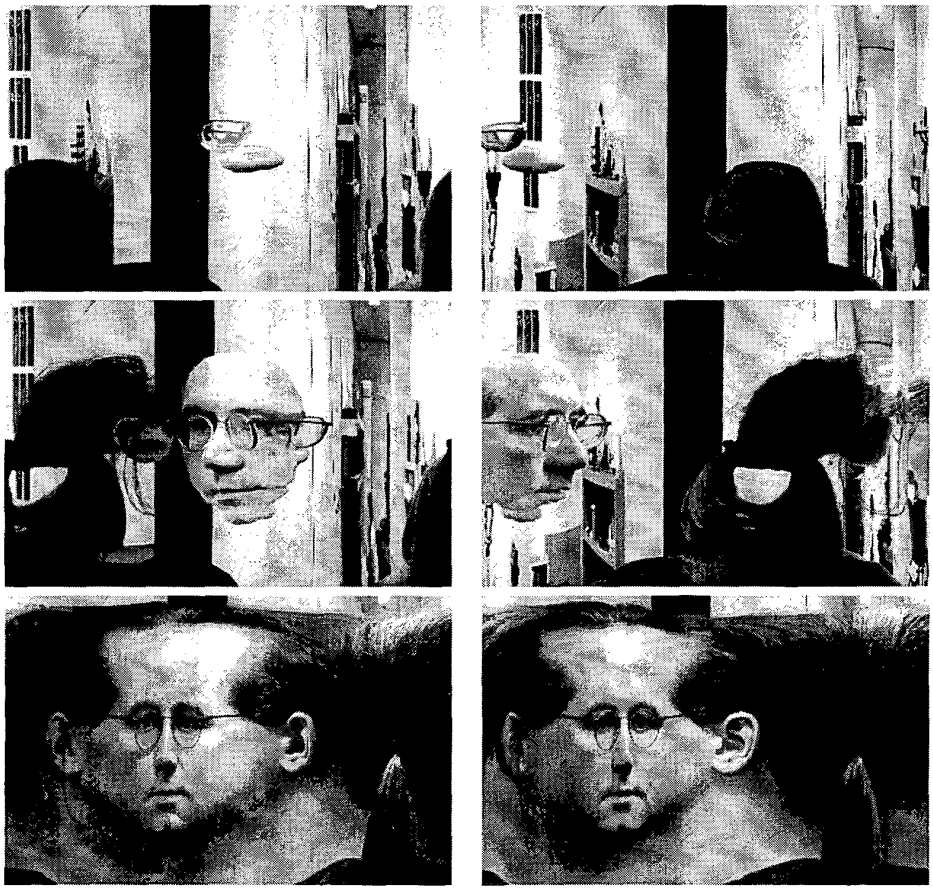


Figure 4: A cyclograph is generated by moving an inward-looking camera on an ellipse and stacking the sequence of images into an x-y-θ volume (a). y-θ slices of the volume form cyclograph images. All of the rays in each cyclograph are tangent to a sheared cylinder (each cross section is an ellipse) (b). A stereo cyclograph pair consists of two views that are both tangent to the same sheared cylinder (b-c), but with tangent rays oriented in opposite directions.

(a)



(b)

Figure 5: Stereo Cyclograph Images. (a) Four of 732 images captured by moving a camera on a circle around a person's head. (b) Cyclograph stereo pairs formed from these input images, corresponding to different slices of the x-y-θ volume.