



The Space of All Stereo Images

STEVEN M. SEITZ AND JIWON KIM

Department of Computer Science and Engineering, University of Washington, Seattle, WA 98195-2350, USA

seitz@cs.washington.edu

jwkim@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. Experimental examples of these images are constructed and used to obtain 3D binocular object reconstructions.

Keywords: stereo, panorama, mosaic, cyclograph, pushbroom, multi-view geometry

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* two-parameter 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 renewed interest in recent years due to the development of practical panoramic sensors (Baker and Nayar, 1999; Nayar and Karmarkar, 2000) and a host of applications in vision and graphics. Furthermore, recent results (Ishiguro et al., 1992) have demonstrated the existence of *stereo panoramas* (Peleg and Ben-Ezra, 1999) 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 (Ishiguro et al., 1992; Shum et al., 1999), 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.

- 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 (2001) 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 a unified mathematics for stereo imaging. The stereo cyclograph is introduced in Section 5 and used experimentally to produce stereo object visualizations and 3D reconstructions.

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* $R(x, y, z, \theta, \phi)$. We represent a light ray as a function R mapping the ray origin (x, y, z) and direction (θ, ϕ) to the set of all points along that ray. 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* (Adelson and Bergen, 1991).

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. We adopt the convention that u varies horizontally and v vertically.

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. The domain \mathcal{D} determines a view's *parameterization*. Cylindrical (McMillan and Bishop, 1995), spherical (Baker and Nayar, 1999), toroidal (Shum et al., 1999), and other non-planar image representations may be represented by choosing a domain with the appropriate topology. 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, encompassing most projection models used in the computer vision literature, including single point perspective, orthographic, weak perspective, and affine models (Koenderink and van Doorn, 1991). In addition, non-perspective mosaic representations fit this definition, including single-perspective panoramas (Baker and Nayar, 1999), multiperspective panoramas (Wood et al, 1997), manifold mosaics (Peleg and Herman, 1997), multiple-center-of-projection images (Rademacher and Bishop, 1998), stereo panoramas (Peleg and Ben-Ezra, 1999), concentric mosaics (Shum and He, 1999), omnivergent images (Shum et al., 1999), and 360×360 panoramas (Nayar and Karmarkar, 2000). 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.

The following terminology and notation will be useful. We use the notation \mathcal{D}^u to denote the restriction of the domain \mathcal{D} to a column u in a view, i.e., $\mathcal{D}^u = \{(a, b) \in \mathcal{D} \mid a = u\}$. Similarly, \mathcal{D}^v denotes the restriction of \mathcal{D} to row v . The shorthand R denotes the set of points along a ray whose parameters are unspecified. Accordingly, we write $\mathbf{X} \in R$ when a 3D point \mathbf{X} lies along a ray R . In this paper, rays are defined to be *open*, i.e., $(x, y, z) \notin R(x, y, z, \theta, \phi)$. This definition is natural in the context of scene visibility—a camera

cannot see its own optical center (x, y, z) . 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, ray, open line segment, 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} .

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 , written $\mathbf{X} \in FOV(V)$ (with some abuse of notation), if \mathbf{X} lies along some ray in $FOV(V)$.

2.2. Multiperspective Imaging

Central (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 (Wood et al., 1997, Rademacher and Bishop, 1998; Shum and He, 1999) and stereo matching (Ishiguro et al., 1992; Shum et al., 1999), 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 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.

Virtually all image representations exhibit some form of continuity, i.e., nearby points in the scene project to nearby points in the image. More formally, 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. While most stereo views of interest exhibit both *u-* and *v-continuity*, our definition of stereo pair requires only the former, allowing for the distribution of rays to vary arbitrarily between rows.

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 stereo pair,³ but yet do not provide parallax cues that can be used to infer depth. Accordingly, we define the stereo viewable 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. Generalized Stereo

It is well-known that any pair of perspective views V_1 and V_2 can be converted into a stereo pair by reparameterization of the respective view domains \mathcal{D}_1 and \mathcal{D}_2 . While typically performed using homographies (Loop and Zhang, 1999), other rectification transforms may be used to avoid certain types of singularities (Roy et al., 1997). Our definition of *stereo pair* is easily generalized to include any pair of views that can be converted into a stereo pair by some rectification transformation:

Two views $V_1 : \mathcal{D}_1 \rightarrow \mathcal{P}$ and $V_2 : \mathcal{D}_2 \rightarrow \mathcal{P}$ satisfy the *generalized stereo constraint* if there exist bijections $r_1 : \mathcal{D}'_1 \rightarrow \mathcal{D}_1$ and $r_2 : \mathcal{D}'_2 \rightarrow \mathcal{D}_2$ such that $V_1 \circ r_1$ and $V_2 \circ r_2$ satisfy the stereo constraint.

This definition reduces the problem of characterizing generalized stereo pairs to that of characterizing (non-generalized) stereo pairs. Note that view geometry, i.e., the distribution of rays in space, is invariant with respect to rectification transformations. Rather, rectification affects only the parameterization of the view domain. By ignoring the specifics of the rectification transforms r_1 and r_2 , we may focus on the simpler case of stereo pairs with horizontal geometry, but ensure that all of our results apply also to the generalized case, which includes views such as catadioptrics with curved epipolar lines (Svoboda et al., 1998).

2.5. 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 pencil of epipolar planes, i.e., epipolar geometry applies 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 (McMillan and Bishop, 1995) or spherical image domains (Svoboda et al., 1998). 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). Consequently, any pair of perspective images satisfies the generalized stereo constraint.

2.6. 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. (1992) and Peleg and Ben-Ezra (1999) 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 (Shum et al., 1999), 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 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

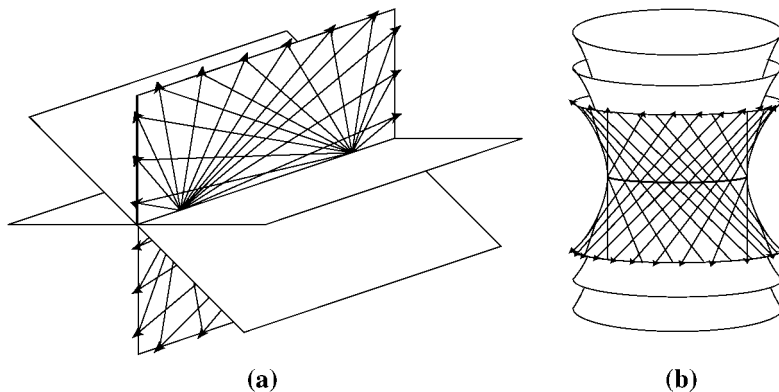


Figure 1. Epipolar geometry. (a) Rays in two perspective views lie on a pencil of epipolar planes. (b) Rays in stereo panoramas and 360×360 panoramas lie on a family of epipolar hyperboloids.

a second hyperboloid S_2 . Note that $S_1 = S_2$, 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.”

Note that Rademacher and Bishop (1998) previously explored ways of defining epipolar geometry for a broad class of multiperspective images called *multiple-center-of-projection* (MCOP) images. Their formulation is quite general, but removes most of the benefits that epipolar geometry provides for stereo processing, in particular: (1) there is generally not a one-to-one correspondence between epipolar curves in two MCOP images, (2) an MCOP image may have a two- rather than a one-dimensional set of epipolar curves, (3) there is no natural extension of epipolar planes to MCOP images, and (4) two MCOP’s cannot always be rectified to produce a stereo pair. We seek a more natural generalization of epipolar geometry to multiperspective images that retains its benefits for stereo processing.

3. A Classification of Stereo Views

What is the space of all stereo views? In this section we derive a complete classification 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 $V_i(u, v)$, for fixed v , sweep out a *ruled surface* S_i . We are interested in the subset of S_1 and S_2 that is stereo viewable. In particular, define the *epipolar surface* for row v to be $S_1^2(v) = S_1 \cap S_2 \cap \mathcal{V}_1^2$. In the case of perspective or any other type of central views, for example, $S_1^2(v)$ is planar for all v .

In the general case, observe that $S_1^2(v)$ contains two families of straight lines, one from V_1 and the other from V_2 . Hence, $S_1^2(v)$ lies on a *doubly-ruled surface* (see Fig. 1). We may therefore characterize the space of stereo 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 (Hilbert and Cohn-Vossen, 1991). While this classical result applies to algebraic surfaces ruled by lines, with some care it can be ex-

tended to the case of arbitrary surfaces ruled by rays or line segments. A proof is presented in Appendix A.

We therefore have the following result:

Stereo Classification Theorem. *V_1 and V_2 are stereo views only if for every row v , $S_1^2(v)$ 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. The generalized stereo constraint in Section 2.4 further broadens 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. 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 hyperbolic 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.

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

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

Generator	Doubly-ruled quadric
Point	Plane
Line	Plane, hyperboloid, hyperbolic paraboloid
Ellipse	Hyperboloid
Parabola	Hyperbolic paraboloid
Hyperbola	Hyperboloid, hyperbolic paraboloid

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 one or more 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 geometry. Table 1 indicates that you could capture the appropriate rays by moving a camera or placing sensors on a line, parabola, or hyperbola. All of the specific quadric view varieties discussed in this paper are generated from a single conic path. It is possible, however, to define quadric views where each row has a different generating path, and these latter views are also subsumed by our framework.

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 panorama* 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 representation which we call

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

Image representation	Generator	Epipolar surfaces
Perspective	Point	Pencil of planes
Stereo panorama (Ishiguro et al., 1992; Peleg and Ben-Ezra, 1999; Shum et al., 1999)	Circle	Hyperboloids
360×360 (Shum et al., 1999; Nayar and Karmarkar, 2000)	Circle	Hyperboloids
Spherical omnivergent (Shum et al., 1999)	Sphere	Pencil of planes
Pushbroom panorama	Line	Pencil of planes
Stereo cyclograph	Ellipse	Hyperboloids
Parabolic panorama	Parabola	Hyperbolic paraboloids

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). This image representation is a generalization of pushbroom cameras (Gupta and Hartley, 1997) that relaxes the assumption that the sensors be oriented perpendicular to the direction of motion and allows for a 360 degree field of view about the axis of translation.

Any two pushbroom panoramas I_i and I_j from the same generating path form a stereo pair with rays lying on a pencil of epipolar planes, as shown in Fig. 2(a). These images are a hybrid between orthographic and perspective images, since rays within each row are parallel but every column of rays meets at a point. They also have the interesting property that disparity is a linear function of depth, in contrast to perspective stereo images. Note that pushbroom panoramas are *not* the same as EPI images (Bolles et al., 1987), although their construction is similar. An example use of pushbroom panoramas is to create a stereo panorama of an entire city street, as shown in Fig. 3.

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). An example of such an image, which we call a *parabolic panorama* is shown in Fig. 4. Two parabolic panoramas that share the same set of hyperbolic paraboloids form a stereo pair. While of theoretical interest, this variety of stereo

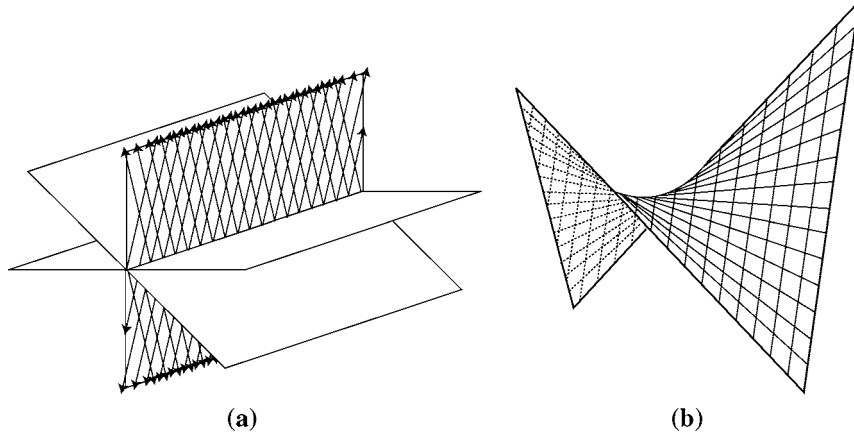


Figure 2. Epipolar geometry of a pushbroom panorama (a) and a parabolic panorama (b). For any row of a pushbroom panorama stereo pair, the pixels from both images lie in a plane. Rays from any row of a parabolic panorama stereo pair trace out a hyperbolic paraboloid.

pair has 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. Towards this end, we are motivated by the following two goals:

1. Unification: operations such as projection and triangulation should be the same for all classes of stereo pairs. In contrast, researchers previously derived specialized equations for each new type of stereo image variety.
2. Generalization: the approach should naturally generalize the case of perspective images, i.e., perspective images and operations should arise naturally as a special case of the new framework.

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 homo-

geneous row vectors $\mathbf{P} = [a \ b \ c \ d]$. Point \mathbf{X} is on plane \mathbf{P} if and only if

$$\mathbf{P}\mathbf{X} = 0 \quad (1)$$

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

$$\mathcal{Q} = \{\mathbf{X} \mid \mathbf{X}^T \mathbf{Q} \mathbf{X} = 0\} \quad (2)$$

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 \mathcal{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 all of the stereo image varieties considered in this paper fit into this category.

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

$$\mathbf{X}^T \mathbf{Q}_u \mathbf{X} = 0 \quad (3)$$

$$\mathbf{X}^T \mathbf{Q}_v \mathbf{X} = 0 \quad (4)$$

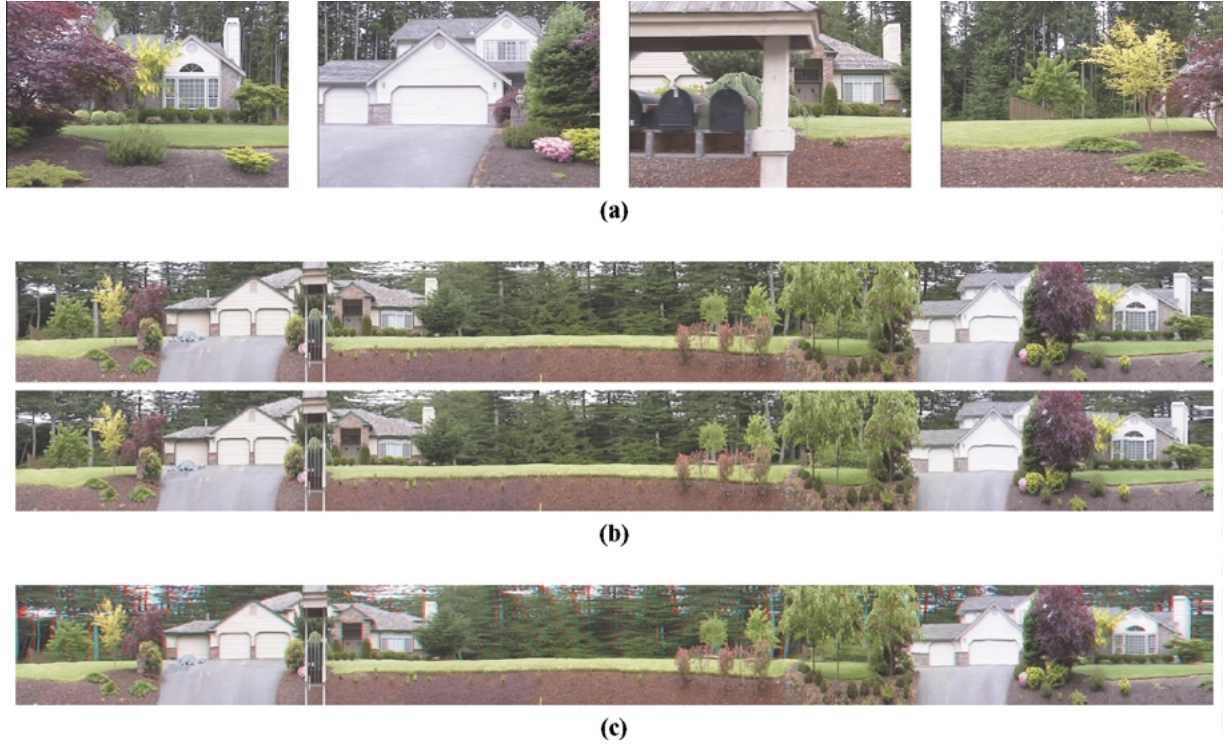


Figure 3. Pushbroom panorama generated from a car driving along a street, with the camera pointing out of the side window. (a) Selected input images. (b) Pushbroom stereo pairs with a baseline of 6, i.e., formed from two columns spaced 6 pixels apart. (c) Red-blue anaglyph generated by taking the blue and green channels from the first image and the red channel from the second image. It should be viewed with red-blue glasses with red over the left eye.

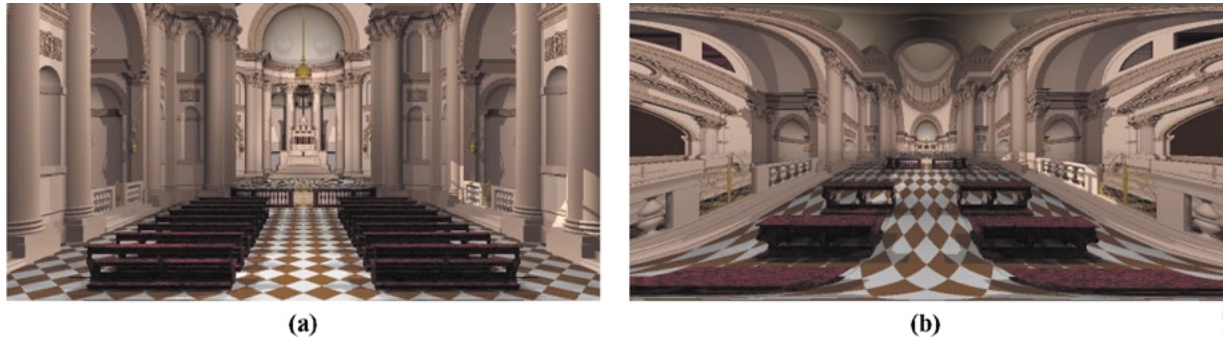


Figure 4. Perspective view (a) and parabolic panorama (b) of a synthetic church scene. (b) was generated by capturing viewing rays along a parabolic camera path. Note that vertical lines remain vertical, but lines are generally not preserved.

where \mathbf{Q}_u is a matrix representing the quadric corresponding to Q_u , and similarly for \mathbf{Q}_v and Q_v . In addition, any number of *field of view* constraints of the following form may be enforced:

$$\mathbf{X}^T \mathbf{Q}_i \mathbf{X} \geq 0 \quad (5)$$

The solutions to Eqs. (3–5) correspond to the set of all points $\mathbf{X} \in FOV(V)$. For the special case that Q_u is a plane for every value of u , Eq. (3) may be more conveniently written:

$$\mathbf{P}_u \mathbf{X} = 0 \quad (6)$$

and similarly for Eqs. (4) and (5).

For example, consider the case of a planar perspective view V . The perspective projection equations may be written: $\mathbf{u} = \mathbf{\Pi}\mathbf{X}$, 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 i th row of $\mathbf{\Pi}$, we can rewrite these equations as follows

$$(\mathbf{\Pi}_1 - u\mathbf{\Pi}_3)\mathbf{X} = 0 \quad (7)$$

$$(\mathbf{\Pi}_2 - v\mathbf{\Pi}_3)\mathbf{X} = 0 \quad (8)$$

plus field of view constraints

$$(\mathbf{\Pi}_1 - u_{\min}\mathbf{\Pi}_3)\mathbf{X} \geq 0 \quad (9)$$

$$-(\mathbf{\Pi}_1 - u_{\max}\mathbf{\Pi}_3)\mathbf{X} \geq 0 \quad (10)$$

$$(\mathbf{\Pi}_2 - v_{\min}\mathbf{\Pi}_3)\mathbf{X} \geq 0 \quad (11)$$

$$-(\mathbf{\Pi}_2 - v_{\max}\mathbf{\Pi}_3)\mathbf{X} \geq 0 \quad (12)$$

Equations (7–12) represent the ray $V(u, v)$ as the intersection of two planes and the viewing frustum, as shown in Fig. 5. $\mathbf{\Pi}_1 - u\mathbf{\Pi}_3$ is the vector corresponding to the plane passing through the camera center and column u of the image plane. Similarly, $\mathbf{\Pi}_2 - v\mathbf{\Pi}_3$ represents the plane from the camera center through row v of the image plane.

As another example, consider the case of a 360×360 panorama (Nayar and Karmarkar, 2000). As originally described in (Shum et al., 1999), 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

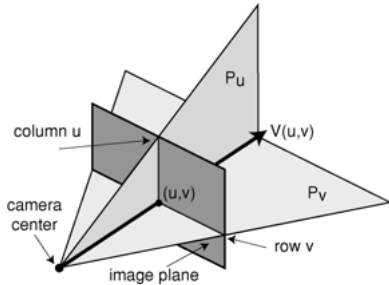


Figure 5. Bi-quadric image representation. A perspective view is represented by two families of planes (quadrics, more generally) \mathbf{P}_u and \mathbf{P}_v and field of view constraints. Each ray in the field of view is specified by the intersection of two of these planes.

follows

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

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

$$\begin{aligned} [-\sin(u) \quad \cos(u) \quad 0 \quad 0] \mathbf{X} &\geq 0 \\ [0 \quad 0 \quad 1 \quad 0] \mathbf{X} &\geq 0 \end{aligned} \quad (15)$$

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. (3) and (4) 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 solutions of u and v are also linear.

In the case of a 360×360 panorama, for example, substituting Eqs. (13) and (14) into Eqs. (6) and (4) yields

$$u = \pm \arccos \left(\frac{x \pm y\sqrt{x^2 + y^2 - 1}}{x^2 + y^2} \right) \quad (16)$$

$$v = \frac{x^2 + y^2 - 1}{z^2} \quad (17)$$

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} :

$$\mathbf{X}^T \mathbf{Q}_{u_1} \mathbf{X} = 0 \quad (18)$$

$$\mathbf{X}^T \mathbf{Q}_{u_2} \mathbf{X} = 0 \quad (19)$$

$$\mathbf{X}^T \mathbf{Q}_v \mathbf{X} = 0 \quad (20)$$

When \mathcal{Q}_{u_1} and \mathcal{Q}_{u_2} are planar, Eqs. (18) and (19) may be replaced with $\mathbf{P}_{u_1} \mathbf{X} = 0$ and $\mathbf{P}_{u_2} \mathbf{X} = 0$, respectively.

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. (3) and (14) into Eqs. (18–20). After applying the field of view constraints (Eqs. (15)) and simplifying, the following expression for \mathbf{X} is obtained:

$$\mathbf{X} = \left[\frac{\sin(u_2) - \sin(u_1)}{\sin(u_2 - u_1)} \quad \frac{\cos(u_1) - \cos(u_2)}{\sin(u_2 - u_1)} \quad \frac{1}{\sqrt{v}} \tan\left(\frac{u_1 - u_2}{2}\right) \quad 1 \right]^T \quad (21)$$

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 (rectified) stereo views:

$$[u_1 \ v_1 \ 1] \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = 0 \quad (22)$$

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. Traditional cyclographs are produced by photographing a rotating object through a narrow slit placed in front of a length of moving film, a technique that dates back to the late 19th century (Davidhazy, 1987).

5.1. Generating Stereo Cyclographs

Using the concepts developed in this paper, it can be seen that the cyclograph construction generates a bi-

quadric image, where the rows of the image sweep out hyperboloids and the columns define planes. In a traditional cyclograph image, all of these planes intersect at the axis of rotation and the quadrics degenerate to (singly-ruled) cones. If instead the slit is moved off-axis, the hyperboloids are not degenerate and are doubly-ruled. In addition, the planes corresponding to each column are all tangent to a cylinder, as shown in Fig. 6(b). Consequently, a stereo pair may be formed. Furthermore, our construction can be generalized to allow the strip to move on any ellipse, rather than a circle. The quadric representation of cyclographs, together with equations for projection and triangulation, are derived in Appendix B.

In order to create stereo cyclograph images of a person’s head, we placed a video camera on a rotating platform and captured a sequence of images while the camera rotated in a circle around the subject. The camera was pointed inward towards the center of rotation. The entire image sequence 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. (1987).⁶ The input images each correspond to an x - y slice through the volume for a fixed value of θ . Cyclograph images correspond to y - θ slices through the volume, as shown in Fig. 6(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.

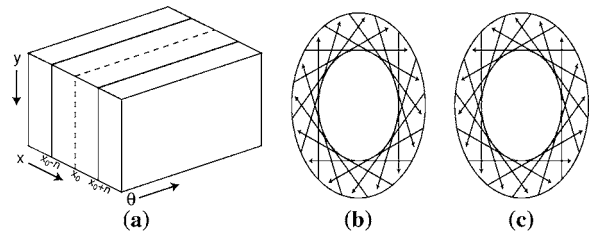


Figure 6. A cyclograph is generated by moving an inward-looking camera on an ellipse and stacking the sequence of images into an x - y - θ volume (a). y - θ slices of the volume form cyclograph images. All of the rays in each cyclograph are tangent to a sheared cylinder (each cross section is an ellipse) (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.



Figure 7. 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. (c) Anaglyph viewable with red-blue stereo glasses (red over left eye).

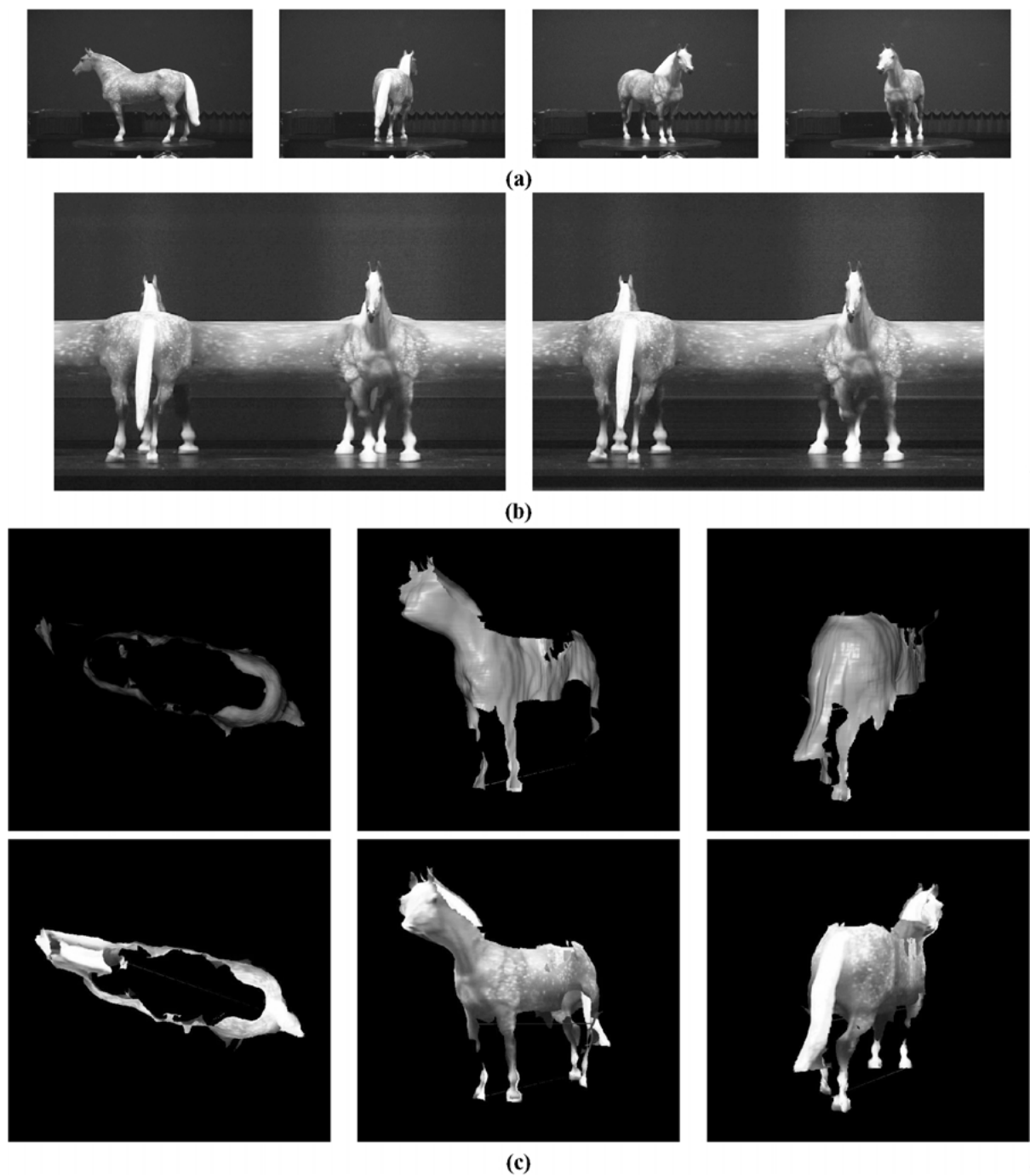


Figure 8. Cyclograph-based reconstruction of a toy horse. (a) Input images. (b) Cyclograph stereo pairs with baseline = 6. (c) Screen shots of reconstructed 3D mesh, shaded (top) and textured (bottom).

Figure 7 shows 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. 7(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 (second row), until the head is

fully visible (third row), creating an interesting effect. The bottom pair is an anaglyph that is viewable with red-blue stereo glasses, with red over the left eye.

Like traditional cyclographs, stereo cyclographs provide a simple and effective means for rendering the appearance of 3D objects for print or display. While traditional cyclographs represent only object appearance, however, stereo cyclographs allow the 3D structure of the object to be retained in the visualization. For instance, surface relief patterns are preserved and readily visualized in the stereo cyclograph images. Since the effective baseline can be varied at display time by choosing different y - θ slices, the stereo pair is easily recalibrated to the eye separation of any observer.

5.2. Shape From Cyclographs

In addition to visualization, cyclographs also have potential benefits for the purposes of shape reconstruction. For instance, the images in the bottom row of Fig. 7(b) could be provided as input to almost any stereo correspondence algorithm, yielding a nearly complete 3D head model from a single image pair. The stereo baseline could be controlled to best suit the stereo algorithm, potentially on a scanline-by-scanline basis. The entire set of cyclographs could be used for structure-from-motion or multi-baseline stereo processing, perhaps using new algorithms that exploit the unique structure of cyclograph sequences.

In this section, we demonstrate the use of stereo matching techniques to reconstruct a 3D object model from a single cyclograph image pair. Cyclographs were generated by placing an object on top of a computer-controlled rotary turntable, and imaging the rotating object from a camera fixed at a certain distance from the table. A cyclograph stereo pair for a fixed baseline was chosen and used as input to a stereo matching technique. We used an (unaltered) implementation of Zitnick and Kanade's technique (Zitnick and Kanade, 2000).

For the results shown here, the background portion of the cyclograph was manually specified and excluded from the reconstruction. Also excluded are pixels that matched with low confidence and points that lie beyond a certain distance from the center of rotation.

Figure 8 shows stereo cyclographs of a toy horse and the resulting binocular stereo reconstructions. An interesting property of cyclograph stereo pairs is that scene points in the cyclograph that are imaged in front of the rotation axis, i.e., between the axis and the cam-

era, have positive disparity, whereas points behind the rotation axis have negative disparity. For instance, in the cyclograph stereo pair of the horse, the front side of the head has positive disparity whereas the back side has negative disparity. This phenomenon is analogous to vergent stereo pairs, in which disparity is positive in front of the vergence point and negative behind it.

The horse reconstruction demonstrates the ability to obtain a true "object model" from a single stereo pair. In contrast, obtaining a model of this sort using traditional stereo images requires stitching together several stereo reconstructions acquired from different viewpoints. The advantage of using cyclographs is that a very reasonable object model may be obtained by running the stereo matching procedure only once, greatly simplifying reconstruction. Note that a reconstruction obtained in this manner may still contain holes—in particular, the top of the horse does not appear in the cyclograph and is thus omitted from the reconstruction. Note also that the head and neck of the horse lack geometric detail in comparison to the rest of the body. This lack of detail is primarily due to the fact that the neck is imaged at an oblique angle in this particular cyclograph pair. In addition, the head moves more quickly in the input sequence since it is closer to the camera path, thus resulting in a lower sampling rate in the cyclographs. As can be seen in Fig. 8(c), however, the shape of the horse is very well reconstructed.

6. Conclusion

A theory of stereo image formation was presented that provides a complete classification of all possible stereo views. A main result is that only three varieties of epipolar surfaces exist, corresponding to rays lying on doubly-ruled quadrics. A second contribution is 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, the *stereo cyclograph*, was experimentally demonstrated and offers compelling advantages for visualization and 3D reconstruction tasks.

While this paper addresses several research questions relating to stereo images, other avenues for future

work in this area remain open. In particular, our definition of stereo image could be considered over-restrictive in certain respects. In particular, we have not considered cases in which multiple viewing rays map to the same pixel, as is the case with lens systems. Extending the theory to cope with multiple-valued views would be desirable and could potentially lead to more stereo varieties.

In addition, it is possible to fuse some image pairs in which the parallax is not perfectly horizontal. Allowing such deviations could also lead to a broader class of images that give a stereo effect, and new *approximate* 3-view and N-view epipolar geometries may be possible (Shum and Szeliski, 1999). Lastly, our analysis was independent of image irradiance information. Taking into account image irradiance might identify images that *appear* to have horizontal parallax, due to special surface texture or shading.

Further exploration of the space of stereo views will be necessary to determine which varieties are useful in practice. The topics of optics and multiperspective camera design are also important areas of future work in which promising initial progress has been made (Nayar and Karmarkar, 2000; Peleg et al, 2000).

Appendix

A. Proof of the Stereo Classification Theorem

We now derive the proof of the Stereo Classification Theorem presented in Section 3. The essence of the proof is to demonstrate that the result holds locally, i.e., that every interior point on $S_1^2(v)$ has a finite neighborhood that lies on a plane, hyperboloid, or hyperbolic paraboloid. Denote as (\mathcal{A}) the interior of a surface \mathcal{A} , i.e., the surface minus its boundary.

Suppose that V_1 and V_2 are a stereo pair of views. Let \mathbf{X} be a point on $(S_1^2(v))$ and (u_i, v) the projection of \mathbf{X} in V_i for $i = 1, 2$. Let L_i be the connected component of $V_i(u_i, v) \cap (S_1^2(v))$ containing \mathbf{X} , as shown in Fig. 9. By *u-continuity*, the projection of L_2 in V_1 spans an interval l_1 of row v that “straddles” u_1 . We say an interval l straddles u if there exists an open set $o \subseteq \Re$ such that $u \in o \subset l$. Observe that the same condition is true if we replace \mathbf{X} with any point on L_1 , i.e., any point $\mathbf{X}' \in L_1$ determines a line segment $L_2' \subset (S_1^2(v))$ whose projection straddles u_1 . Let L_2' and L_2'' be two such line segments. Define

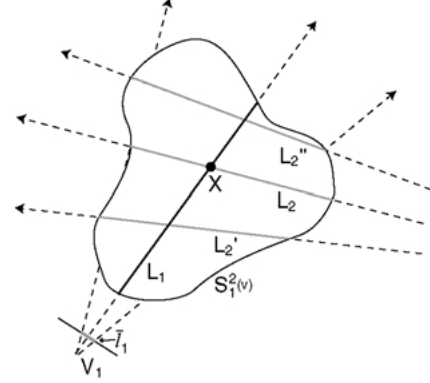


Figure 9. Proof of the Stereo Classification Theorem. There is an open interval \bar{l}_1 of row v of V_1 such that every ray in this interval intersects all three line segments L_2 , L_2' , and L_2'' . This condition implies that the rays in \bar{l}_1 must lie on a plane, hyperboloid, or hyperbolic paraboloid.

l_1' and l_1'' to be the respective projections of L_2' and L_2'' in V_1 .

Consider $\bar{l}_1 = l_1 \cap l_1' \cap l_1''$. By construction, \bar{l}_1 is an interval that straddles u_1 . Furthermore, $V_1(u, v)$ intersects L_2 , L_2' , and L_2'' for all $u \in \bar{l}_1$. We therefore have an infinite family of rays $\mathcal{N} = \{V_1(u, v) \mid u \in \bar{l}_1\}$ that all intersect three distinct lines, through L_2 , L_2' , and L_2'' . Interestingly, this condition uniquely defines the surface in the neighborhood of \mathbf{X} (Hilbert and Cohn-Vossen, 1991). In particular, if two of L_2 , L_2' , and L_2'' are co-planar then \mathcal{N} must lie on a plane. If the line segments are not co-planar but do lie on parallel planes, \mathcal{N} lies on a hyperbolic paraboloid. Otherwise, \mathcal{N} lies on a hyperboloid.

We have shown that every $\mathbf{X} \in (S_1^2(v))$ has an open neighborhood that lies on a quadric surface, more specifically, a plane, hyperboloid, or hyperbolic paraboloid. It follows that every connected component of $(S_1^2(v))$ lies on the same quadric.

In particular, if $(S_1^2(v))$ is itself connected, a condition satisfied by most representations of interest including all varieties described in this paper, it lies entirely on a plane, hyperboloid, or hyperbolic paraboloid.

B. Cyclograph Representation and Equations

This appendix presents the quadric representation of a stereo cyclograph and derives the equations for projection and triangulation. We present the equations for the case of a circular generator, but they may be generalized to the case of cyclographs generated by elliptical

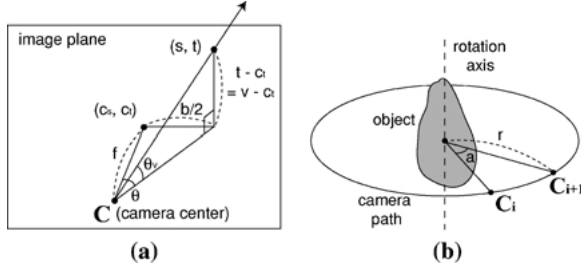


Figure 10. (a) Geometry of a perspective view (note: the dotted line denotes the length of the corresponding line segment). (b) Camera moving along a circular path around the object. C_i denotes the position of the camera center at frame i .

camera paths. We introduce the following terminology (see Fig. 10 for graphic illustrations):

- Coordinate systems
 - Input view: (s, t) whose origin is bottom-left corner of the image plane
 - World space: (x, y, z) whose origin is the intersection of the rotation axis and the camera motion plane
 - Cyclograph: (u, v) whose origin is bottom-left corner of the image
- a : amount of rotation between frames in radians
- b : baseline in pixels, i.e., distance between the two columns from which the stereo cyclograph pair is generated
- f : camera's focal length in pixels
- c_t : t coordinate of the camera's principal point
- r : radius of the circular camera path
- $\theta = \arctan(\frac{b/2}{f})$
- $\theta_v = \arctan(\frac{t-c_t}{\sqrt{f^2 + \frac{b^2}{4}}}) = \arctan(\frac{v-c_v}{\sqrt{f^2 + \frac{b^2}{4}}})$

The quadric representation of a stereo cyclograph is specified as follows:

$$\mathbf{P}_u^+ = [\sin(ua + \theta) \quad -\cos(ua + \theta) \quad 0 \quad -r \sin \theta] \quad (23)$$

$$\mathbf{P}_u^- = [\sin(ua - \theta) \quad -\cos(ua - \theta) \quad 0 \quad r \sin \theta] \quad (24)$$

$$\mathbf{Q}_v = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{\tan^2 \theta_v} & \frac{r \cos \theta}{\tan \theta_v} \\ 0 & 0 & \frac{r \cos \theta}{\tan \theta_v} & -r^2 \end{bmatrix} \quad (25)$$

with the following field of view constraints:

$$[-\cos(ua + \theta) \quad -\sin(ua + \theta) \quad 0 \quad r \cos \theta] \mathbf{X} \geq 0 \quad (26)$$

$$[-\cos(ua - \theta) \quad -\sin(ua - \theta) \quad 0 \quad r \cos \theta] \mathbf{X} \geq 0 \quad (27)$$

$$\mathbf{X}^T \begin{bmatrix} 1 & 0 & 0 & -r \cos(ua) \\ 0 & 1 & 0 & -r \sin(ua) \\ 0 & 0 & -\frac{1}{\tan^2 \theta_v} & 0 \\ -r \cos(ua) & -r \sin(ua) & 0 & r^2 \end{bmatrix} \times \mathbf{X} = 0 \quad (28)$$

Equations (23) and (26) correspond to the right image, and Eqs. (24) and (27) correspond to the left image of the stereo pair. The field of view constraints, Eqs. (26) and (27), specify that this view captures only the “forward” rays in front of the camera. The last field of view constraint Eq. (28) enforces the viewing ray to pass through the camera center, and is used to determine a unique solution for triangulation.

B.1. Derivation. Rays from a column of a cyclograph rule out a plane. Figure 11(a) shows the plane ruled out by rays of column u of the left image of the stereo pair, which is generated by taking a column to the right of the principle point from each input frame. The right image is analogously formed by taking from each frame a column offset by the same distance to the left of the principal point. Although this might seem counter-intuitive, it is similar to the perspective vergent stereo where the same scene point is imaged by a right column of the left image and a left column of the right image, as illustrated by Fig. 11(b).

Equations (23) and (26) represent the plane ruled out by rays of column u . Since the angle of the plane from the x - z plane is $ua \pm \theta$ ($ua - \theta$ for the left image of stereo pair as shown in Fig. 11(a), and $ua + \theta$ for the right image), and it passes through the camera center $\mathbf{C} = (r \cos(ua), r \sin(ua), 0)$, its equation is given as follows:

$$y - r \sin(ua) = \tan(ua \pm \theta) \cdot (x - r \cos(ua))$$

Rearranging the above equation yields

$$\begin{aligned} & x \sin(ua \pm \theta) - y \cos(ua \pm \theta) + r(\sin(ua) \\ & \quad \times \cos(ua \pm \theta) - \cos(ua) \sin(ua \pm \theta)) = 0 \\ & \Leftrightarrow x \sin(ua \pm \theta) - y \cos(ua \pm \theta) \mp r \sin \theta = 0 \\ & \Leftrightarrow \mathbf{P}_u \mathbf{X} = 0 \end{aligned}$$

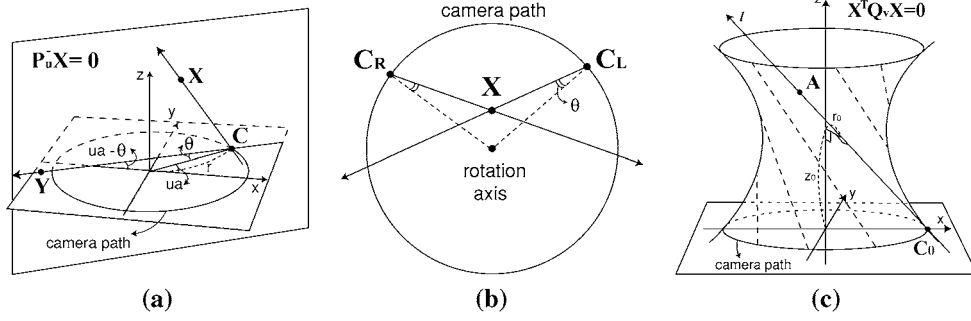


Figure 11. (a) Plane ruled out by rays of column u . (b) Top-down view of (a), showing both planes that correspond to the left and right image of the stereo pair. C_L and C_R denotes the camera centers of the left and right image, respectively. (c) Hyperboloid ruled out by rays of row v .

As shown in Fig. 11(c), rays from a row of a cyclograph rule out a hyperboloid. In general, a hyperboloid can be expressed by an equation of the form $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$. Since the camera path is circular and the center of the hyperboloid is translated along the z axis, the hyperboloid equation can be written as

$$x^2 + y^2 - \frac{(z - z_0)^2}{k} = r_0^2$$

As illustrated in Fig. 11(c), z_0 corresponds to the z coordinate of the translated center of the hyperboloid, and r_0 corresponds to the minimum radius of the x - y cross section of the hyperboloid occurring at $z = z_0$.

We need to solve for three coefficients k , z_0 and r_0 . As shown in Fig. 10(a), z^2 increases by $\tan^2 \theta_v$ whenever $x^2 + y^2$ increases by 1. Therefore, $k = \tan^2 \theta_v$.

To compute z_0 and r_0 , without loss of generality, consider a viewing ray l which passes through a camera center at $C_0 = (r, 0, 0)$. From the illustration in Fig. 10(a), an arbitrary point A on l can be expressed as $(-ft + r, \frac{b}{2} \cdot t, (v - c_t)t)$. We need to find t that minimizes $x^2 + y^2$, i.e., the distance from the z axis.

$$\begin{aligned} x^2 + y^2 &= (-ft + r)^2 + \frac{b^2}{4} \cdot t^2 \\ &= \left(f^2 + \frac{b^2}{4}\right) \left(t - \frac{rf}{f^2 + \frac{b^2}{4}}\right)^2 + \frac{r^2 \cdot \frac{b^2}{4}}{f^2 + \frac{b^2}{4}} \\ &= \left(f^2 + \frac{b^2}{4}\right) \left(t - \frac{r \cos \theta}{\sqrt{f^2 + \frac{b^2}{4}}}\right)^2 + r^2 \sin^2 \theta \end{aligned}$$

The minimum value of $x^2 + y^2$ is $r^2 \sin^2 \theta$, which corresponds to r_0^2 . Also, since the minimum occurs at $t = \frac{r \cos \theta}{\sqrt{f^2 + \frac{b^2}{4}}}$, z_0 is obtained by substituting t into the z

coordinate of A .

$$z_0 = (v - c_t) \cdot \frac{r \cos \theta}{\sqrt{f^2 + \frac{b^2}{4}}} = r \cos \theta \tan \theta_v$$

Therefore, the equation of the hyperboloid finally reduces to

$$\begin{aligned} x^2 + y^2 - \frac{1}{\tan^2 \theta_v} (z - r \cos \theta \tan \theta_v)^2 - r^2 \sin^2 \theta &= 0 \\ \Leftrightarrow X^T Q_v X &= 0 \end{aligned}$$

Equations (26) and (27) are field of view constraints stating that this view contains only the subset of rays that point toward the inside of the circle. As shown in Fig. 11(a), for X to lie on the “forward” portion of the viewing ray, the dot product between vector \vec{CX} and \vec{CY} must be nonnegative, where Y is a point on the line which is the intersection of the plane $P_u X = 0$ and the plane containing the camera path.

$$\vec{CX} \cdot \vec{CY} \geq 0$$

$$\begin{aligned} \Leftrightarrow (x - r \cos(ua), y - r \sin(ua), z) \\ \cdot (-\cos(ua \pm \theta), -\sin(ua \pm \theta), 0) &\geq 0 \\ \Leftrightarrow -x \cos(ua \pm \theta) - y \sin(ua \pm \theta) \\ + r(\cos(ua) \cos(ua \pm \theta) \\ + \sin(ua) \sin(ua \pm \theta)) &\geq 0 \\ \Leftrightarrow -x \cos(ua \pm \theta) - y \sin(ua \pm \theta) + r \cos \theta &\geq 0 \\ \Leftrightarrow [-\cos(ua \pm \theta) \quad -\sin(ua \pm \theta) \quad 0 \quad r \cos \theta] \\ \times X &\geq 0 \end{aligned}$$

Equation (28) is a second field of view constraint enforcing X to lie on the viewing ray from the camera

center \mathbf{C} . Since the angle of $\overrightarrow{\mathbf{CX}}$ with respect to $x - y$ plane is θ_v ,

$$\begin{aligned} \tan \theta_v &= \frac{z}{\sqrt{(x - r \cos(ua))^2 + (y - r \sin(ua))^2}} \\ \Leftrightarrow (x - r \cos(ua))^2 + (y - r \sin(ua))^2 \\ &\quad - \frac{z^2}{\tan^2 \theta_v} = 0 \end{aligned}$$

which yields Eq. (28). This constraint is necessary to obtain a unique solution for triangulation, as described in the following subsection.

B.2. Projection and Triangulation. Without loss of generality, we assume that the camera rotates around the object in the counter-clockwise direction, and the cyclograph is generated by concatenating columns from left to right.

The 2D projection (u, v) of a 3D point \mathbf{X} into the left image of the stereo pair is computed by solving $\mathbf{P}_u^- \mathbf{X} = 0$ and $\mathbf{X}^T \mathbf{Q}_v \mathbf{X} = 0$, which yields:

$$u = \frac{1}{a} \cdot \left(\arccos \left(\frac{-yr \sin \theta \pm x \cdot \sqrt{x^2 + y^2 - r^2 \sin^2 \theta}}{x^2 + y^2} \right) + \theta \right) \quad (29)$$

$$\mathbf{X} = \begin{bmatrix} \frac{r \sin \theta \cos \left(\frac{u_1 + u_2}{2} a \right)}{\sin \left(\frac{u_1 - u_2}{2} a + \theta \right)} & \frac{r \sin \theta \sin \left(\frac{u_1 + u_2}{2} a \right)}{\sin \left(\frac{u_1 - u_2}{2} a + \theta \right)} & \frac{r \tan \theta_v \sin \left(\frac{u_1 - u_2}{2} a \right)}{\sin \left(\frac{u_1 - u_2}{2} a + \theta \right)} & 1 \end{bmatrix}^T \quad (36)$$

$$v = \frac{z \cdot \sqrt{f^2 + \frac{b^2}{4}}}{r \cos \theta \pm \sqrt{x^2 + y^2 - r^2 \sin^2 \theta}} + c_t \quad (30)$$

Analogously, the projection into the right image of the stereo pair is obtained by solving the equations $\mathbf{P}_u^+ \mathbf{X} = 0$ and $\mathbf{X}^T \mathbf{Q}_v \mathbf{X} = 0$, which yields:

$$u = \frac{1}{a} \cdot \left(\arccos \left(\frac{yr \sin \theta \pm x \cdot \sqrt{x^2 + y^2 - r^2 \sin^2 \theta}}{x^2 + y^2} \right) - \theta \right) \quad (31)$$

$$v = \frac{z \cdot \sqrt{f^2 + \frac{b^2}{4}}}{r \cos \theta \pm \sqrt{x^2 + y^2 - r^2 \sin^2 \theta}} + c_t \quad (32)$$

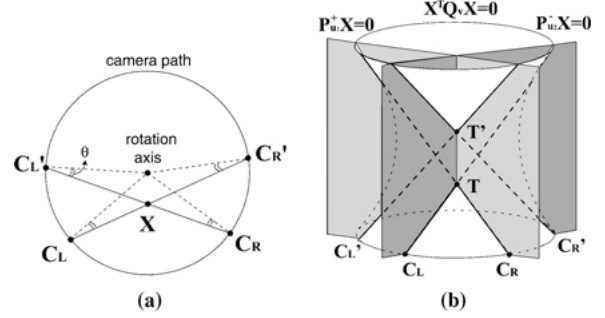


Figure 12. (a) A scene point \mathbf{X} projects twice into the same cyclograph. (b) Solving the triangulation equation produces two solutions \mathbf{T} and \mathbf{T}' , and \mathbf{T}' is a false solution that is eliminated by enforcing the second field of view constraint.

In fact, Eqs. (30) and (32) turn out to be the same, implying that identical solutions for v are obtained for the left and right image.

The triangulation of (u_1, v) and (u_2, v) corresponds to solving the following equations for \mathbf{X} :

$$\mathbf{P}_{u_1}^+ \mathbf{X} = 0 \quad (33)$$

$$\mathbf{P}_{u_2}^- \mathbf{X} = 0 \quad (34)$$

$$\mathbf{X}^T \mathbf{Q}_v \mathbf{X} = 0 \quad (35)$$

Enforcing the field of view constraints, Eqs. (26), (27) and (28), yields the following expression for \mathbf{X} :

The second field of view constraint (Eq. (28)) eliminates a false solution that does not originate from the camera centers given by u_1 and u_2 , as illustrated by Fig. 12(b). Each plane intersects the hyperboloid along two lines, and the intersection of these lines produces two triangulation points. However, only one of these two points is the true solution that corresponds to the intersection of two rays that originate from the correct camera centers \mathbf{C}_L and \mathbf{C}_R .

Acknowledgments

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 stereo

matching algorithm for the horse reconstruction was provided by Charles Zitnick and is available via the web: (<http://www.cs.cmu.edu/~clz/stereo.html>). The authors would like to thank Shai Avidan for helpful discussions in the early stages of this work, and also the anonymous reviewers of this paper for several useful comments and suggestions. The support of Intel Corporation, Microsoft Corporation, and the National Science Foundation under grants IIS-9984672 and IIS-0049095 is gratefully acknowledged.

Notes

1. For the purposes of this paper, relativistic effects due to the curvature of space are ignored.
2. In RGB color space, \mathcal{E} may be represented by \mathbb{R}^3 . More generally, \mathcal{E} can be a function space representing the light spectrum.
3. Note that all rays of a perspective view converge at a common point, the camera center \mathbf{C} . However, the rays do not actually intersect at \mathbf{C} , since the rays are defined to be open and do not contain the camera center. This condition is quite natural and conforms to the intuition that a ray contains only points that are potentially visible to the camera, i.e., points that are in front of \mathbf{C} .
4. More specifically, “horizontal” means in the direction of a row on the image plane of the first view.
5. Cyclographs are also referred to as *periphotographs* or *rollout photographs*.
6. The *EPI* images introduced by Bolles et al. (1987) correspond to x - t slices of an image volume obtained by moving a camera along a line.

References

- Adelson, E.H. and Bergen, J.R. 1991. *The Plenoptic Function and the Elements of Early Vision*. MIT Press: Cambridge, MA.
- Baker, S. and Nayar, S. 1999. A theory of single-viewpoint catadioptric image formation. *Int. J. of Computer Vision*, 35(2):175–196.
- Bolles, R.C., Baker, H.H., and Marimont, D.H. 1987. Epipolar-plane image analysis: An approach to determining structure from motion. *Int. J. of Computer Vision*, 1(1):7–55.
- Davidhazy, A. 1987. Principles of peripheral photography. *Industrial Photography*. Available at <http://www.rit.edu/~andpph/text-peripheral-basics.html>.
- Gupta, R. and Hartley, R.I. 1997. Linear pushbroom cameras. *IEEE Trans. Pattern Analysis and Machine Intell.*, 19(9):963–975.
- Hilbert, D. and Cohn-Vossen, S. 1991. *Geometry and the Imagination*. AMS Chelsea: Providence, RI.
- Ishiguro, H., Yamamoto, M., and Tsuji, S. 1992. Omni-directional stereo. *PAMI*, 14:257–262.
- Koenderink, J.J. and van Doorn, A.J. 1991. Affine structure from motion. *J. Opt. Soc. Am. A*, 8:377–385.
- Loop, C. and Zhang, Z. 1999. Computing rectifying homographies for stereo vision. In *Proc. Computer Vision and Pattern Recognition Conf.*
- McMillan, L. and Bishop, G. 1995. Plenoptic modeling: An image-based rendering system. In *Proc. SIGGRAPH 95*, pp. 39–46.
- Nayar, S.K. and Karmarkar, A. 2000. 360×360 mosaics. In *Proc. Computer Vision and Pattern Recognition Conf.*, pp. 388–395.
- Pajdla, T. 2001. Epipolar geometry of some non-classical cameras. In *Proceedings of Computer Vision Winter Workshop*. Slovenian Pattern Recognition Society, pp. 223–233.
- Peleg, S. and Ben-Ezra, M. 1999. Stereo panorama with a single camera. In *Proc. Computer Vision and Pattern Recognition Conf.*, pp. 395–401.
- Peleg, S. and Herman, J. 1997. Panoramic mosaics by manifold projection. In *Proc. Computer Vision and Pattern Recognition Conf.*, pp. 338–343.
- Peleg, S., Pritch, Y., and Ben-Ezra, M. 2000. Cameras for stereo panoramic imaging. In *Proc. Computer Vision and Pattern Recognition Conf.*, pp. 208–214.
- Rademacher, P. and Bishop, G. 1998. Multiple-center-of-projection images. In *Proceedings of SIGGRAPH 98*, pp. 199–206.
- Roy, S., Meunier, J., and Cox, I. 1997. Cylindrical rectification to minimize epipolar distortion. In *Proc. Computer Vision and Pattern Recognition Conf.*, pp. 393–399.
- Shum, H.-Y. and He, L.-W. 1999. Rendering with concentric mosaics. In *Proceedings of SIGGRAPH 99*, pp. 299–306.
- Shum, H.-Y., Kalai, A., and Seitz, S.M. 1999. Omnivergent stereo. In *Proc. 7th Int. Conf. on Computer Vision*, pp. 22–29.
- Shum, H.-Y. and Szeliski, R. 1999. Stereo reconstruction from multiperspective panoramas. In *Proc. Seventh Int. Conf. on Computer Vision*, pp. 14–21.
- Svoboda, T., Pajdla, T., and Hlavac, V. 1998. Epipolar geometry for panoramic cameras. In *Proc. 5th Eur. Conf. on Computer Vision*, pp. 218–232.
- Wood, D.N., Finkelstein, A., Hughes, J.F., Thayer, C.E., and Salesin, D.H. 1997. Multiperspective panoramas for cel animation. In *Proc. SIGGRAPH 97*, pp. 243–250.
- Zitnick, C.L. and Kanade, T. 2000. A cooperative algorithm for stereo matching and occlusion detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(7).