Epsilon Stereo Pairs

Yuanyuan Ding Jingyi Yu
Department of Computer and Information Sciences
University of Delaware
Newark, DE 19716, USA
{ding,yu}@cis.udel.edu

Abstract

Human stereo vision works by fusing a pair of perspective images with a purely horizontal parallax. Recent developments suggest that very few varieties of multiperspective stereo pairs exist. In this paper, we introduce a new stereo model, which we call epsilon stereo pairs, for fusing a broader class of multiperspective images. An epsilon stereo pair consists of two images with a slight vertical parallax. We show many multiperspective camera pairs that do not satisfy the stereo constraint can still form epsilon stereo pairs. We then introduce a new ray-space warping algorithm to minimize stereo inconsistencies in an epsilon pair using multiperspective collineations. This makes epsilon stereo model a promising tool for synthesizing close-to-stereo fusions from many non-stereo pairs.

1 Introduction

A stereo pair consists of two images with a purely horizontal parallax. They provide important depth cues that are amenable to processing using computer vision algorithms [1, 4]. The classic pinhole and orthographic camera models have long served as the workhorse in capturing stereo pairs. However, recent developments have suggested that cameras that do not adhere to the pinhole structure may also form stereo pairs. For instance, a stereo panorama fuses two pushbroom images to synthesize a 360 degree depth perception [8]. We refer to these images as *multiperspective stereo pairs*.

Seitz [10] has classified all possible stereo pairs in terms of their epipolar geometry. Pajdla [6] independently obtained a similar result. They have shown that the epipolar geometry, if it exists, has to be a double ruled surfaces. Therefore, very few varieties of multiperspective stereo pairs exist. In this paper, we introduce a new framework, which we call epsilon stereo pairs, to model a much broader class of multiperspective images.

An epsilon stereo pair (ε-pair) consists of two images with a slight vertical parallax. We show many previous multiperspective cameras form valid epsilon stereo pairs even though they do not satisfy the stereo constraint. We use ε-pairs to model the recently proposed General Linear Cameras (GLC) [11]. We show both vertical and horizontal parallaxes can be directly derived from the GLC intrinsics.

We then present a new ray-space warping algorithm to minimize stereo inconsistencies in an ε-pair. Our approach is based on the multiperspective collineation theory [12], which describes the transformation between the images of a camera due to changes in

sampling and image plane selection. We show a proper collineation can effectively reduce stereo inconsistency in an ε-pair. Finally, we develop an automatic algorithm to find an optimal collineation via non-linear optimizations. This makes the epsilon stereo model a promising tool for synthesizing close-to-stereo fusions from many non-stereo pairs.

2 Previous Work

Classic pinhole cameras collect rays passing through a single point. Because of its simplicity, any oblique pair of pinhole images can be warped to have a purely horizontal parallax via projective transformations (homography) [4]. Recently, several researchers have proposed alternative multiperspective camera models, which capture rays originating from different points in space. These multiperspective cameras include pushbroom cameras [3], which collect rays along parallel planes from points swept along a linear trajectory, two-slit cameras [13], which collect all rays passing through two lines, and oblique cameras [7], in which each pair of rays are oblique.

Despite their incongruity of view, some multiperspective cameras can still form valid stereo pairs. Peleg et. al. [8] stitched the same column of images from a rotating pinhole camera to form a circular pushbroom. They then fused two oblique circular pushbrooms to synthesize a stereo panorama. Feldman et. al. proved that a pair of cross-slit cameras can have valid epipolar geometry if they share a slit or the slits intersect in four pairwise distinct points [2]. Seitz [10] and Pajdla [6] independently classified the space of all possible stereo pairs in terms of the epipolar geometry. Their work suggests that only three varieties of epipolar geometry exist: planes, hyperboloids, and hyperbolic-paraboloids, all corresponding to double ruled surfaces.

This paper focuses on how to fuse multiperspective camera pairs that do not have valid epipolar geometry. Such pairs may consist of two different cross-slit cameras, a cross-slit and a pushbroom, or two arbitrary multiperspective cameras. To simplify our analysis, we use the recently proposed General Linear Cameras (GLC) [11] to uniformly model these cameras.

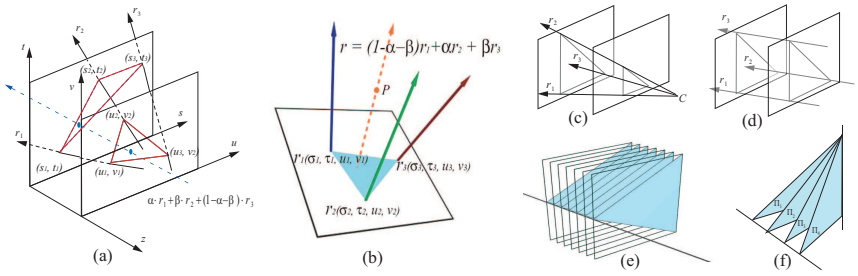

Figure 1: General Linear Cameras. (a) A GLC collects radiance along all possible affine combination of three rays. The rays are parameterized by their intersections with two parallel planes. (b) The GLC projection maps every 3D point P to a ray. The GLC model unifies many previous cameras, including the pinhole (c), the orthographic (d), the pushbroom (e), and the cross-slit (d).

2.1 General Linear Cameras

In the GLC framework, every ray is parameterized by its intersections with the two parallel planes, where $[s,t]$ is the intersection with the first and $[u,v]$ the second, as shown

in Figure 1(a). This parametrization is often called a two-plane parametrization (2PP) [5, 11]. Alternatively, we can reparameterize each ray by substituting $\sigma = s - u$ and $\tau = t - v$. In this paper, we will use this $[\sigma, \tau, u, v]$ parametrization to simplify our analysis. We also assume the default uv plane is at $z = 0$ and st plane at $z = 1$. Thus $[\sigma, \tau, 1]$ represents the direction of the ray.

A GLC is defined as the affine combination of three rays:

$$GLC = \{r : r = \alpha \cdot [\sigma_1, \tau_1, u_1, v_1] + \beta \cdot [\sigma_2, \tau_2, u_2, v_2] + (1 - \alpha - \beta) \cdot [\sigma_3, \tau_3, u_3, v_3], \forall \alpha, \beta\} \tag{1}$$

Many well-known multiperspective cameras, such as push-broom, cross-slit, linear oblique cameras are GLCs.

We further simply the GLC model by choosing three specific rays that have $[u, v]$ coordinates as $[0,0]$, $[1,0]$, and $[0,1]$ to form a canonical GLC:

$$r[\sigma, \tau, u, v] = (1 - \alpha - \beta) \cdot [\sigma_1, \tau_1, 0, 0] + \alpha \cdot [\sigma_2, \tau_2, 1, 0] + \beta \cdot [\sigma_3, \tau_3, 0, 1] \tag{2}$$

It is easy to see that $\alpha = u$ and $\beta = v$. Therefore, every pixel $[u, v]$ maps to a unique ray in the GLC.

Given a 3D point $P[x, y, z]$, the GLC projection maps P to pixel $[u, v]$ (see [12]) as:

$$u = \frac{\begin{vmatrix} z\sigma_1 & z\tau_1 & 1 \\ x & y & 1 \\ z\sigma_3 & 1 + z\tau_3 & 1 \end{vmatrix}}{Az^2 + Bz + C}, \quad v = \frac{\begin{vmatrix} z\sigma_1 & z\tau_1 & 1 \\ 1 + z\sigma_2 & z\tau_2 & 1 \\ x & y & 1 \end{vmatrix}}{Az^2 + Bz + C} \tag{3}$$

The denominator corresponds to the characteristic equation of the GLC:

$$Az^2 + Bz + C = 0 \tag{4}$$

where

$$A = \begin{vmatrix} \sigma_1 & \tau_1 & 1 \\ \sigma_2 & \tau_2 & 1 \\ \sigma_3 & \tau_3 & 1 \end{vmatrix}, B = \begin{vmatrix} \sigma_1 & v_1 & 1 \\ \sigma_2 & v_2 & 1 \\ \sigma_3 & v_3 & 1 \end{vmatrix} - \begin{vmatrix} \tau_1 & u_1 & 1 \\ \tau_2 & u_2 & 1 \\ \tau_3 & u_3 & 1 \end{vmatrix}, C = \begin{vmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \\ u_3 & v_3 & 1 \end{vmatrix} \tag{5}$$

The root $z_i (i = 1, 2)$ to equation (4) corresponds to a slit (line) on plane $z = z_i$ that all rays in the GLC will simultaneously pass through [11]. For instance, the cross-slit characteristic equation has two distinct roots since all rays simultaneously pass through two slits whereas an oblique camera has no solution.

Notice points on these slits will project to infinity on the image plane. To avoid this singularity, we assume the scene geometry $S(S^x, S^y, S^z)$ lie on the positive side of the image plane and is bounded by volume $[x_{min}, y_{min}, z_{min}] \times [x_{max}, y_{max}, z_{max}]$, where $z_{min} > max(z_1, z_2, 0)$.

3 Epsilon Stereo Pairs

A stereo pair consists of two images with a pure horizontal parallax, i.e., for every 3D point P, its images $[u, v]$ and $[u', v']$ in the two cameras must satisfy $v = v'$. We introduce a new stereo model, which we call epsilon stereo pairs, for fusing camera pairs that do not satisfy the stereo constraint.

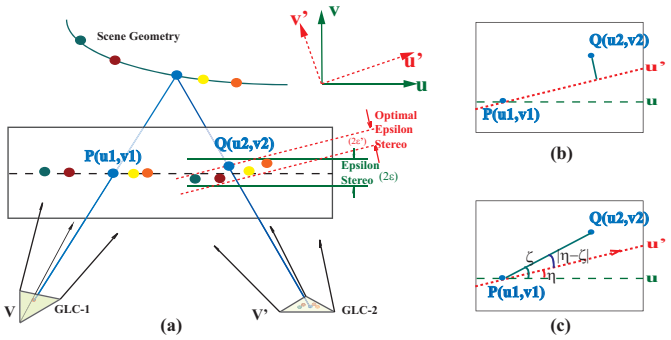

Figure 2: (a) An epsilon stereo pair consists of two images with a mostly horizontal parallax and a slight (ε vertical parallax. The vertical parallax is measured using a distance metric (b) or an angular metric (c). We can change the horizontal direction $\vec{d}$ to reduce ε.

We say that two views V and V' form an *epsilon stereo pair* if the following property holds:

The rays $V(u,v)$ and $V'(u',v')$ intersect only if $|v-v'| \leq \varepsilon$. The classical stereo constraint is a special case of the epsilon stereo model when $\varepsilon = 0$.

Any two such views are referred to as an ε-pair. In our analysis, we make the same assumption as [10], i.e., all views are *u*- and *v-continuous*. Furthermore, we only consider scene geometry visible in both views.

Physically, an ε-pair represents two views with a mostly horizontal parallax and with a slight (ε) vertical parallax. In an ε-pair, all pixels on a row in V correspond to pixels lying inside the ε band around the same row in V', as shown in Figure 2. When ε is small enough, our visual system can potentially fuse the two images as if they were stereo pairs, as shown in Figure 4.

3.1 Translation GLC Pairs

We show that any two GLCs form a valid ε-Pair. We first consider the pairs in which the second GLC is a translation of the first GLC along the image plane. Assume the first GLC is specified in its canonical form $[\sigma_1,\tau_1,0,0]$, $[\sigma_2,\tau_2,0,0]$, and $[\sigma_3,\tau_3,0,0]$, and the translation is specified as $[-t_x,-t_y,0]$.

The projection of a point $\dot{P}(x,y,z)$ in the first GLC can be computed using Equation (3). To project P to the second GLC, we can simply translate $\dot{P}$ by $[t_x,t_y,0]$. The vertical and the horizontal parallax of $\dot{P}$ can be computed as:

$$\Delta v = v - v' = -\frac{\begin{vmatrix} z\sigma_1 & z\tau_1 & 1 \\ 1+z\sigma_2 & z\tau_2 & 1 \\ t_x & t_y & 0 \end{vmatrix}}{Az^2+Bz+C} = \frac{z(t_y(\sigma_1-\sigma_2)+t_x(\tau_2-\tau_1))-t_y}{Az^2+Bz+C} \quad (6)$$

$$\Delta u = u - u' = -\frac{\begin{vmatrix} z\sigma_1 & z\tau_1 & 1 \\ t_x & t_y & 0 \\ z\sigma_3 & 1+z\tau_3 & 1 \end{vmatrix}}{Az^2+Bz+C} = \frac{z(t_y(\sigma_3-\sigma_1)+t_x(\tau_1-\tau_3))-t_x}{Az^2+Bz+C} \quad (7)$$

Notice, the x and y terms diminish in Δv and Δu. Thus, for a translation GLC pair, the vertical parallax Δv is a quadratic rational function of z. To show they form a valid epsilon stereo pair, we only need to prove $\Delta v(z)$ is bounded. We sketch the proof as follows.

Without loss of generality, we assume the characteristic equation (4) has coefficient $A > 0$ and has two roots z_1 and z_2 (if $A < 0$, we can analyze -$\Delta v(z)$ instead). Thus, equation (4) can be rewritten as $Az^2 + Bz + C = A(z - z_1)(z - z_2)$. We then replace it into (6) as:

$$\Delta v(z) = \frac{z(t_y(\sigma_1 - \sigma_2) + t_x(\tau_2 - \tau_1)) - t_y}{A(z - z_1)(z - z_2)} = \frac{\frac{1}{A}(t_y(\sigma_1 - \sigma_2) + t_x(\tau_2 - \tau_1))z - \frac{t_y}{A}}{(z - z_1)(z - z_2)} \tag{8}$$

Recall that the scene depth satisfies $z > max(z_1, z_2, 0)$, therefore, the denominator of (8) is positive and monotonically increasing. Next, we consider the numerator of equation (8).

If $\frac{1}{A}(t_y(\sigma_1 - \sigma_2) + t_x(\tau_2 - \tau_1)) < 0$, then the numerator is monotonically decreasing. In this case, $\Delta v(z)$ is monotonically decreasing, and its extrema can be computed as:

$$\Delta v(z)_{min} = \lim_{z \to \infty} \Delta v(z) = 0, \quad \Delta v(z)_{max} = \Delta v(z_{min}) \tag{9}$$

If $\frac{1}{A}(t_y(\sigma_1 - \sigma_2) + t_x(\tau_2 - \tau_1)) > 0$, we can rewrite the equation (8) as:

$$\Delta v(z) = \frac{\frac{1}{A}(t_y(\sigma_1 - \sigma_2) + t_x(\tau_2 - \tau_1)) - \frac{t_y}{Az}}{(1 - \frac{z_1}{z})(z - z_2)} \tag{10}$$

For $z > max(z_1, z_2)$, the denominator remains positive and monotonically increasing. For the numerator, when $t_y < 0$, it is monotonically decreasing. Thus, $\Delta v(z)$ is a monotonically decreasing function and its bound can be derived similarly to the derivation in (9). When $t_y > 0$, the numerator is a monotonically increasing. However it is bounded with maximum $\frac{1}{A}(t_y(\sigma_1 - \sigma_2) + t_x(\tau_2 - \tau_1))$ at $z = \infty$ and minimum at $z = z_{min}$. Thus, $\Delta v(z)$ is bounded, although it may not be monotonic w.r.t. z.

We have shown that Δv, in general, is bounded. In fact, in many cases, Δv is monotonically decreasing with respect to z. A similar derivation applies to Δu. Notice, the horizontal parallax, Δu, provides an important depth cue of the scene. The monotonicity in Δu is particularly important in stereo fusion. Our analysis indicates translation GLC pairs are not only ε-pairs but also well-suited for multiperspective stereo fusion if ε is small.

3.2 General GLC Pairs

Next, we consider a pair of two different GLCs. Assume the two GLCs are specified as $[\sigma_1, \tau_1, \sigma_2, \tau_2, \sigma_3, \tau_3]$, and $[\sigma_1', \tau_1', \sigma_2', \tau_2', \sigma_3', \tau_3']$. We have

$$\Delta v(x, y, z) = v - v' = \frac{\begin{vmatrix} z\sigma_1 & z\tau_1 & 1 \\ 1 + z\sigma_2 & z\tau_2 & 1 \\ x & y & 1 \end{vmatrix}}{Az^2 + Bz + C} - \frac{\begin{vmatrix} z\sigma_1' & z\tau_1' & 1 \\ 1 + z\sigma_2' & z\tau_2' & 1 \\ x & y & 1 \end{vmatrix}}{A'z^2 + B'z + C'} \tag{11}$$

where $(x, y, z) \in [x_{min}, y_{min}, z_{min}] \times [x_{max}, y_{max}, z_{max}]$

Since scene geometry lies beyond the slits of both GLCs, the denominator $Az^2 + Bz + C$ and $A'z^2 + B'z + C'$ cannot be zero. Furthermore Δv is linear in x and y and is bounded by $[x_{min}, x_{max}] \times [y_{min}, y_{max}]$. Therefore, we only need to consider Δv w.r.t z.

Recall that $v(z)$ and $v'(z)$ are both rational functions in z with quadratic numerators and denominators. Let z_1, z_2 be the two roots of the corresponding GLC characteristic equation for $v(z)$. We can rewrite the projection equation (3) as:

$$v(z) = \frac{\frac{1}{A}(\sigma_1\tau_2 - \sigma_2\tau_1) + \frac{1}{A}(y(\sigma_2 - \sigma_1) + x(\tau_1 - \tau_2) - \tau_1)\frac{1}{z} + \frac{y}{A}\frac{1}{z^2}}{(1 - \frac{z_1}{z})(1 - \frac{z_2}{z})} \tag{12}$$

Let $\kappa = \frac{1}{z}$,since scene points satisfy $z > max(z_1, z_2, z_{min})$, we have $\kappa \in (0, \frac{1}{max(z_1,z_2,z_{min})})$, and equation (12) transforms to:

$$v(\kappa) = \frac{\frac{1}{A}(\sigma_1\tau_2 - \sigma_2\tau_1) + \frac{1}{A}(y(\sigma_2 - \sigma_1) + x(\tau_1 - \tau_2) - \tau_1)\kappa^1 + \frac{y}{A}\kappa^2}{(1 - z_1\kappa)(1 - z_2\kappa)}$$

Thus, we have $(1 - z_1\kappa)(1 - z_2\kappa) \in ((1 - \frac{z_1}{max(z_1,z_2,z_{min})})(1 - \frac{z_2}{max(z_1,z_2,z_{min})}), 1)$. This proves $v(z)$ is bounded w.r.t to z. A similar proof applies to $v'(z)$, and hence, Δv is also bounded.

Notice, our analysis only applies to GLCs whose characteristic equations are quadratic, such as cross-slit and linear oblique cameras. The characteristic equation may degenerate to linear in the case of pushbroom cameras. However, in that case, the GLC projection is linear and rational, and a similar derivation can be applied.

3.3 Reducing Vertical Parallax

The vertical parallax in an ε-pair can be further reduced by using a different horizontal direction, as shown in Figure 2(a). Here, we present a simple and intuitive algorithm for finding the optimal horizontal direction for a GLC ε-pair. We assume that the two GLCs share the same image plane at $z = 0$. Our goal is to estimate the optimal horizontal direction $\vec{d}$ so that the average vertical parallax ε is minimal.

Assume the scene geometry is known, then, for each 3D point we compute projections in the ε pair as $P[u_1, v_1]$ and $Q[u_2, v_2]$ using equation (3). To measure the vertical parallax, we can compute the distance from Q to the scanline that passes through P with direction $\vec{d}$, as shown in Figure 2(b).

Assume the angle between $\vec{d}$ and the u-axis is η, this distance can be computed as:

$$d_{PQ} = |sin\eta(u_2 - u_1) - cos\eta(v_2 - v_1)| \tag{13}$$

The optimal η can be obtained by solving the following miminization problem:

$$\eta_{opt} = \min_{\eta} \int_{\Theta} |sin\eta(u_2 - u_1) - cos\eta(v_2 - v_1)|^2 dxdydz \tag{14}$$

Alternatively, we can measure the angular difference between $\vec{PQ}$ and $\vec{d}$ as shown in Figure 2(c). Let ζ denote the angle between $\vec{PQ}$ and u axis, we have:

$$\eta - \zeta = \eta - \arctan(\frac{v_2 - v_1}{u_2 - u_1}) \tag{15}$$

Hence the optimal horizontal direction can be computed as:

$$\eta_{opt} = \min_{\eta} \int_{\Theta} (\eta - \arctan(\frac{v_2 - v_1}{u_2 - u_1}))^2 dxdydz \tag{16}$$

4 Optimal Epsilon Stereo Pairs

When an ε-pair consists of cameras that capture significantly different rays, changing the horizontal direction on the image plane is insufficient for reducing the vertical parallax. In this section, we present a new ray-space warping algorithm based on multiperspective collineations.

4.1 Multiperspective Collineation

A multiperspective collineation describes the transformation between the images of the camera due to changes in sampling and image plane selection. This transformation is analogous to planar collineation (homography) in pinhole cameras.

A collineation can be specified by an image plane $\Pi[\dot{p}, \vec{d}_1, \vec{d}_2]$, where $\dot{p}$ specifies the origin of the plane and $\vec{d}_1$ and $\vec{d}_2$ specify the two spanning directions. For every ray r that has origin $\dot{o}$ and a direction $\vec{l}$, we intersect ray r with Π to compute its pixel coordinate $[i, j]$ as:

$$[o_x, o_y, o_z] + \lambda [l_x, l_y, l_z] \quad = \quad \dot{p} + \vec{d}_1 i + \vec{d}_2 j \tag{17}$$

Solving for i, j, and λ in (17) gives:

$$\begin{aligned} i &= \frac{(-l^z d_2^y + l^y d_2^z)(o^x - p^x) + (l^z d_2^x - l^x d_2^z)(o^y - p^y) + (-l^y d_2^x + l^x d_2^y)(o^z - p^z)}{\gamma} \\ j &= \frac{(l^z d_1^y - l^y d_1^z)(o^x - p^x) + (-l^z d_1^x + l^x d_1^z)(o^y - p^y) + (l^y d_1^x - l^x d_1^y)(o^z - p^z)}{\gamma} \end{aligned} \tag{18}$$

where

$$\gamma = \begin{vmatrix} d_1^x & d_2^x & -l^x \\ d_1^y & d_2^y & -l^y \\ d_1^z & d_2^z & -l^z \end{vmatrix} \tag{19}$$

For a canonical GLC, $[\dot{p}, \vec{d}_1, \vec{d}_2]$ are all linear functions of the pixel coordinate $[u, v]$. Therefore, the GLC collineation is, in general, a quadratic rational function.

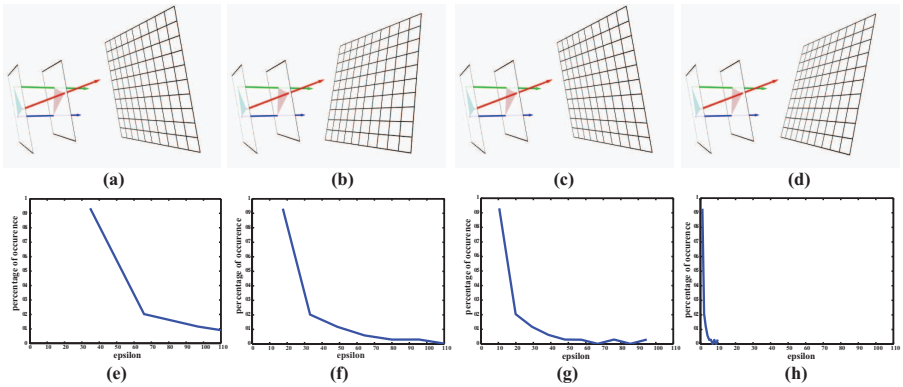

Figure 3: We use multiperspective collineations to reduce the vertical parallax between a pinhole and a cross-slit camera. Top row: We apply different collineations on the cross-slit camera. Bottom row: We plot the corresponding distribution of the vertical parallax. The vertical parallax is significantly reduced using collineation (d).

4.2 Automatic Collineation Optimization

Given an ε-pair, we use collineations to reduce stereo inconsistencies. In Figure 3, we apply different collineations to an ε-pair that consists of a pinhole camera and a cross-slit camera. We also densely sample the scene and plot the vertical parallax distributions

under each collineation using a histogram. The horizontal axis represents the range of vertical parallaxes and the vertical represents the occurrence percentage of the points. As shown in Figure 3(h), the vertical parallax of an ε-pair can be significantly reduced with a proper collineation.

We also present an automatic algorithm to robustly estimate the optimal collineation. Our algorithm starts with features matching between the two images. Recall that each GLC captures a different set of rays and exhibits unique distortions. Thus, matching the feature points between two GLCs can be difficult. In our implementation, we use Scale Invariant Feature Transform (SIFT) to preprocess the images. SIFT robustly handles image distortion and illumination variations and generates transformation-invariant features. We then perform global matching using SSD to find the potential matching pairs. To remove outliers, our algorithm uses RANSAC where we use projective transformations as its precondition.

Given feature correspondences $f_k(i_k, j_k), (k = 1, \ldots, n)$ in I and $f'_k(i'_k, j'_k), (k = 1, \ldots, n)$ in I', we compute the ray representation r'_k for each feature f'_k in I'. Our goal is to find the optimal collineation $[\dot{p}, \vec{d}_1, \vec{d}_2]$ that minimizes the vertical parallax between f_k and $Proj(r_k, \Pi)$, where $Proj(r'_k, \Pi)$ represents the projection of ray r'_k under collineation Π. As shown in equation (18). We formulate this optimization as a least squares problem:

$$\min_{\Pi} \sum_{r'_k \in featureraysinI'} D^2(Proj(r'_k, \Pi), f_k) \tag{20}$$

The distance metric $D^2(Proj(r'_k, \Pi), f_k)$ measures the vertical parallax between f_k and $Proj(r'_k, \Pi)$.

One way to compute D is to calculate the vertical distance between $Proj(r'_k, \Pi)$ and f_k. In Figure 4(b), we apply the vertical distance metric to fuse two cross-slit images. We use blue arrows to mark the displacements between the feature points in the final fusion. Although the vertical parallax is significantly reduced, many features pairs, such as the ones of the middle building, lie very close to each other and fail to provide correct depth cues.

We propose a different distance metric that simultaneously measures the vertical and the horizontal parallax for each feature pair. We estimate the intersection point of the corresponding two rays for each pair and assign an approximated horizontal disparity $disp_k$ that emulates how it would be seen in a perspective stereo pair. We formulate the new distance metric as:

$$D^2(Proj(r'_k), f_k) = (Proj(r'_k)^i - (f_k^i + disp_k))^2 + (Proj(r'_k)^j - f_k^j)^2 \tag{21}$$

Finally we use Levenberg-Marquardt to solve equation (20).

5 Results

We have performed experiments on various GLC pairs using our algorithm. We have modified the PovRay [9] ray tracer to accommodate different GLC types. In Figure 4, we fuse a pinhole image with a cross-slit image of a city scene. These two cameras cannot form epipolar geometry. To better illustrate our results, we color the pinhole image with green and the cross-slit with red and fuse the two images in an anaglyph. We apply

our automatic collineation algorithm to find the optimal collineations. We highlight the displacement between the corresponding features with blue arrows in the final anaglyph. The initial two images exhibit a severe vertical parallax (Figure 4(a)). We then compute the optimal collineation using the vertical distance. As a result, the vertical parallax is significantly reduced. However, many features lie too close to each other and do not provide valid depth cues (Figure 4(b)). We then use the disparity metric (21) to compute the best collineation. The resulting collineation produces a close-to-stereo pair with a minimal vertical parallax and a consistent horizontal parallax (Figure 4(c)).

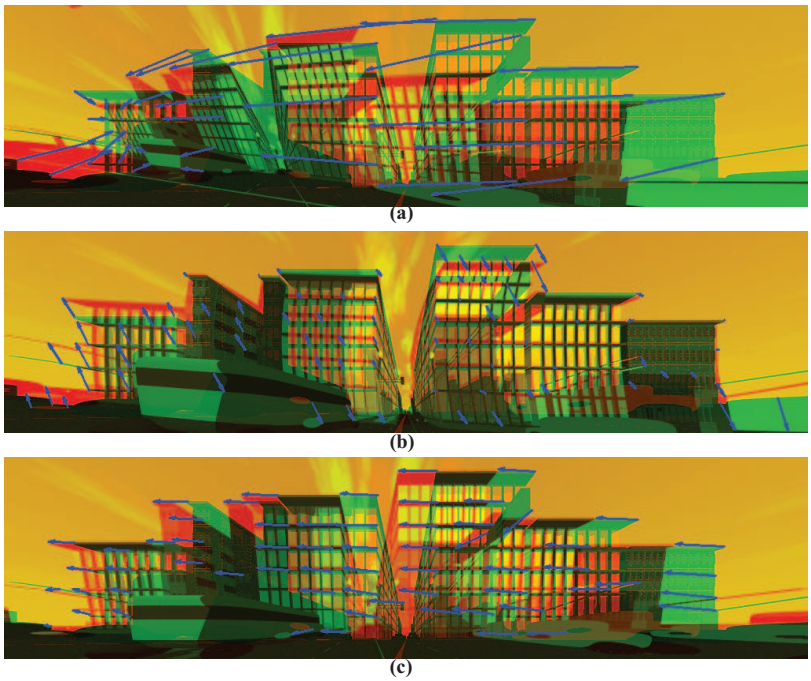

Figure 4: (a) An anaglyph of a pinhole (green) and a cross-slit (yellow). (b) An optimized ε-pair using the vertical parallax metric. Overall, the vertical parallax is reduce. However, the depth cue is also lost. (c) An optimized ε-pair using the disparity metric. The vertical parallax is reduced and the horizontal parallax is preserved.

We have also compared our collineation algorithm with the commonly used projective transformation. In Figure 5, we fuse a pinhole image (Figure 5(a)) and a pushbroom image with additional synthetic radial distortions (Figure 5(b)). Figure 5(c) computes the optimal projective transformation and Figure 5(d) computes the optimal collineation, both using the disparity metric (21). The optimal collineation result is less distorted and is more consistent with the pinhole image. This is because multiperspective collineation describes a much broader class of warping functions than the projective transformation.

6 Conclusion

In this paper, we have introduced a new framework, which we call epsilon stereo pairs, to model images captured by a broader class of multiperspective cameras. An epsilon

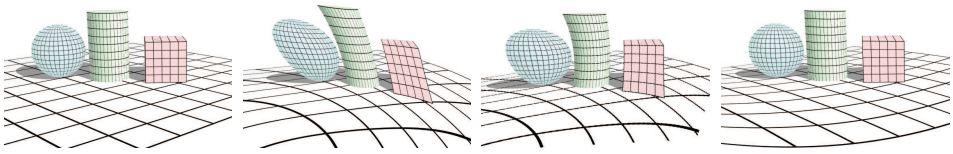

Figure 5: We compare collineation with the projection transformation on an ε-pair of a pinhole camera (a) and a pushbroom camera with synthetic radial distortions (b). (c) shows the the optimal projective transformation result. (d) shows our optimal collineation result.

stereo pair consists of two images that have a slight vertical parallax. We have shown many previous multiperspective images form epsilon stereo pairs, even though they do not satisfy the stereo constraint. We have also proposed a new ray-space warping algorithm based on multiperspective collineations to minimize the vertical parallax between two images. Experiments demonstrate that our epsilon stereo model provides a promising tool for synthesizing close-to-stereo fusions using non-stereo pairs.

Acknowledgement

This work has been supported by the National Science Foundation under grant NSF-MSPA-MCS-0625931.

References

[1] Bolles, R. C., H. H. Baker, and D. H. Marimont: Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion. International Journal of Computer Vision, Vol. 1 (1987).

[2] D. Feldman, T. Pajdla, and D. Weinshall: On the Epipolar Geometry of the Crossed-Slits Projection. In Proc. 9th Int. Conf. on Computer Vision (2003).

[3] R. Gupta and R.I. Hartley: Linear Pushbroom Cameras. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 9 (1997) 963–975.

[4] R.I. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision. Cambridge Univ. Press, 2000.

[5] M. Levoy and P. Hanrahan: Light Field Rendering. Proc. ACM SIGGRAPH '96 31–42.

[6] T. Pajdla: Epipolar Geometry of some non-classical cameras. Proceedings of Computer Vision Winter Workshop, pp. 223-233, Slovenian Pattern Recognition Soceity.

[7] T. Pajdla: Stereo with Oblique Cameras. Int'l J. Computer Vision, vol. 47, nos. 1/2/3 (2002) 161–170.

[8] S. Peleg, M. Ben-Ezra, and Y. Pritch: Omnistereo: Panoramic Stereo Imaging. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 3 (2001) 279–290.

[9] POV-Ray: The Persistence of Vision Raytracer. http://www.povray.org/

[10] S. M. Seitz: The Space of All Stereo Images. Proc. Int'l Conf. Computer Vision '01, vol. I (2001) 26–33.

[11] J. Yu and L. McMillan: General Linear Cameras. Proceedings of 8th European Conference on Computer Vision (2004), Volume 2, pp. 14-27.

[12] J. Yu and L. McMillan: Multiperspective Projection and Collineation. Proceedings of 10th Intl. Conferenc on Computer Vision (2005), pp. 580-587.

[13] A. Zomet, D. Feldman, S. Peleg, and D. Weinshall: Mosaicing New Views: The Crossed-Slits Projection. IEEE Trans. on PAMI (2003) 741–754.