Contour-Based Correspondence for Stereo

Shamez Alibhai and Steven W. Zucker*

Center for Computational Vision and Control
Depts. of Computer Science and Electrical Engineering
Yale University
New Haven, CT U.S.A.
steven.zucker@yale.edu

Abstract. In stereoscopic images, the behavior of a curve in space is related to the appearance of the curve in the left and right image planes. Formally, this relationship is governed by the projective geometry induced by the stereo camera configuration and by the differential structure of the curve in the scene. We propose that the correspondence problem–matching corresponding points in the image planes–can be solved by relating the differential structure in the left and right image planes to the geometry of curves in space. Specifically, the compatibility between two pairs of corresponding points and tangents at those points is related to the local approximation of a space curve using an osculating helix. To guarantee robustness against small changes in the camera parameters, we select a specific osculating helix. A relaxation labeling network demonstrates that the compatibilities can be used to infer the appropriate correspondences in a scene. Examples on which standard approaches fail are demonstrated.

1 Introduction

The objects in our visual environment weave through space in an endless variety of depths, orientations and positions (Figs. 1, 2). Nevertheless, in computer vision the dominant theme has been to develop region-based methods to solve the correspondence problem [6,7,11,13,14,16], or else to focus on long, straight lines[5]. Edge features have been introduced to reduce the complexity of matching, but ambiguities along the epipolar lines (Fig. 1) are dealth with by a distance measure, e.g. similarity of orientation, over the possible edge matches. While this filters the number of potential matches, it does not solve the problem of which edge elements in each figure should correspond.

To resolve these remaining ambiguities, global (heuristic) constraints have been introduced, including the uniqueness constraint [6,13,16]; the ordering constraint (the projections of two objects in an image must have the same left to right ordering as the objects in the scene) [1,6,16]; the smoothness constraint [2, 6,7,13,14], and so on. However, each of these is heuristic, and breaks down for

* Research supported by AFOSR

D. Vernon (Ed.): ECCV 2000, LNCS 1842, pp. 314–330, 2000.

Fig. 1. A stereo pair of twigs. There are many places in the stereo pair where the ordering constraint breaks down. The images were taken using a verged stereo rig and a baseline of 9cm.

the above images; see Fig. 3. It is our goal to develop a stereo correspondence system that can work for scenes such as these.

The occluding contour of an object weaves through space as does the object with which it is associated. The projection of the occluding contour is an edge curve that winds across the image plane. We focus on the relationship between the differential-geometric properties of curves in a scene and the projection of those curves into the left and right image planes. These observations allow us to explore the reasons why conventional stereo systems may fail for natural images and will provide the basis for our new approach.

The distance measures that have been defined on features in the image plane, together with global constraints, have not fully captured the relationships between a space curve in the scene and the behavior of the projected curves in the (left, right) image planes. For example, consider matching edge elements' by orientation along an epipolar line: if the threshold for determining a match were too high (typical values include ± 30 deg [7,16]) then too many possible matches would be admitted; if the threshold were too low, then the system would be restricted to detecting edges that do not traverse through depth. Imposing additional measures, such as the curvature of lines in the image plane, does not solve the problem of ambiguity because, like the measures on orientation, measures on curvature do not take into account the behavior of curves in three space. The structure of the curves in the image plane must derive from the geometry of the space curve.

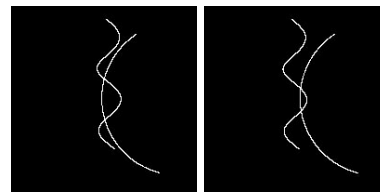

Fig. 2. A synthetic stereo pair consisting of a y-helix and a circular arc.

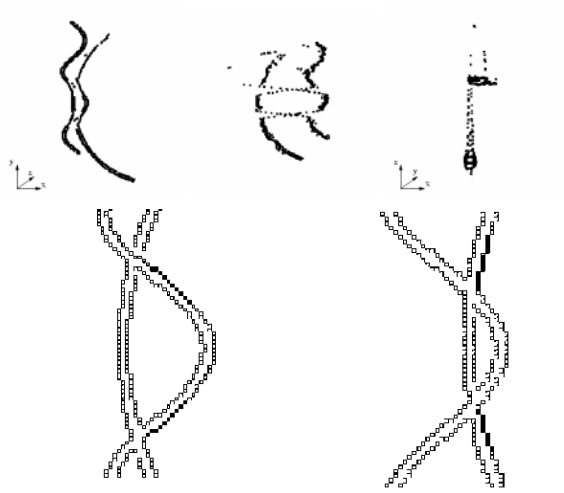

Fig. 3. The false matches obtained using the PMF algorithm of Pollard, Mayhew, and Frisby in the TINA system on the example in Fig. 2. (TOP) Three orthogonal projections of the reconstruction. The rightmost "birds-eye view" illustrates the complete merging of two curves in depth. The incorrect reconstruction is due to the failure of the ordering constraint in the image, which is highlighted in (BOTTOM) zoom. The (filled) darkened pixels in the left and right images correspond in TINA according to the ordering constraint, but are clearly incorrect.

Specifically, we will relate the local approximation of a curve in three space to its projection in the left and right image planes. The behavior of a planar curve can be described in terms of its tangent and curvature. Similarly, a curve in three space can be described by the relationships between its tangent, normal and binormal. As the curve moves across depth planes, there exists a positional disparity between the projection of the curve in the left image and the projection in the right image. However, there also exist higher order disparities, for example disparities in orientation, that occur. It is these types of relationships that can be capitalized upon when solving the correspondence problem. One approach is to relate them directly to surface slant and tilt [12]. Rather than defining a distance measure that compares the positions or orientations of a left/right edge-element pair, we will require that neighboring pairs be locally consistent. That is to say: (i) we shall interpret edge elements as signaling tangents; and (ii) we shall require that there exists a curve in three space whose projection in the left and right image planes is commensurate with the locus of tangent pairs in a neighborhood of the proposed match.

The reduction of the correspondence problem to relations between nearby tangents distributed across (left, right) image pairs opens up a second goal for this research - biologically plausible stereo correspondence algorithms that can run in orientation hypercolumns. However, space limitations preclude us from further developing this goal.

Clearly, the role of projective geometry must not be ignored. One would not want correspondence relations that depended in minute detail on camera parameters and precise camera calibration, otherwise the solution would be too delicate to use in practice. We seek those relationships between the differential structure of a space curve and the differential structure of image curves on which stereo correspondence is based to be invariant to small changes in the camera parameters. We refer to this as the PRINCIPLE OF DIFFERENTIAL INVARIANCE.

2 Differential Formulation of the Compatibility Field

Let $\alpha(s)$ be a curve in $\Re^3$. Assuming measurements of the curvature, torsion and Frenet frame of the curve at $\alpha(0)$, we obtain a local approximation of the curve by taking the third order Taylor expansion of α at $s = 0$. Using the Frenet equations and keeping only the dominant term in each component, we obtain the FRENET APPROXIMATION of α at $s = 0$ (see Fig. 4):

$$\hat{\alpha}(s) = \alpha(0) + sT_0 + \kappa_0 \frac{s^2}{2} N_0 + \kappa_0 \tau_0 \frac{s^3}{6} B_0 \tag{1}$$

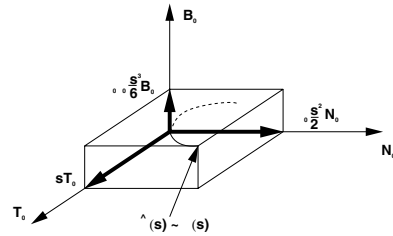

Fig. 4. The Frenet approximation to a curve in $\Re^3$ at a point $\alpha(0)$. Without loss of generality, the point $\alpha(0)$ is assumed to be coincident with the origin.

Another approximation can be derived for the case of a planar curve. For the moment let α be a unit speed planar curve with $\kappa > 0$. There is one and only one OSCULATING CIRCLE ψ which approximates α near $\alpha(s)$ up to second-order in the Frenet sense. This approximation was used in defining the CO-CIRCULARITY CONSTRAINT for determining consistency between nearby edge elements [15]. Our technical goal in this section is to generalize co-circularity to stereo. The following abstraction is useful:

TRANSPORT PROBLEM Given a unit speed curve α, determine a unit vector field, $\boldsymbol{u}$, on α such that $\alpha'(s) \cdot \boldsymbol{u}(s) = \text{const}$.

For any unit speed curve, α, the vector field $\boldsymbol{u} = T$ can be defined such that the pointwise relation $\alpha' \cdot \boldsymbol{u} = 1$ holds. If a curve is reconstructed from an image, then the reconstructed curve must satisfy this property. However, this

assumes–unrealistically!– that α was known around the point $\alpha(0)$. The basis for co-circularity is the observation that α can be locally approximated at 0 via the osculating circle, ψ: if the tangent at the position $\psi(\delta s)$ is transported along the osculating circle from $\psi(0)$ to $\psi(\delta s)$, it should approximate the tangent at $\alpha(\delta s)$ to first order. This TRANSPORT CONSTRAINT was used to define compatibilities for a relaxation labeling process.

For the stereo correspondence problem, two edge maps are given (one for the left camera and one for the right); each of these will be consistent (in the sense that they satisfy the monocular co-circularity transport constraint); our goal now is to make them consistent with a local approximation to the space curve from which they project. The generalization of the osculating circle for space curves requires an osculating helix. To begin, recall that a unit speed circular helix is a space curve $h(s)$ where $h(s) = \left(a\cos\left(\frac{s}{c}\right), a\sin\left(\frac{s}{c}\right), \frac{bs}{c}\right)$ where $c = \sqrt{a^2+b^2}$.
OSCULATING HELIX Let α be a unit speed space curve with $\kappa > 0$. There is one and only one unit speed circular helix, $\tilde{h}$, that locally approximates α at s in the sense of Frenet: $\tilde{h}(0) = \alpha(s); \boldsymbol{T}_{\tilde{h}}(0) = \boldsymbol{T}_{\alpha}(s); \boldsymbol{N}_{\tilde{h}}(0) = \boldsymbol{N}_{\alpha}(s); \kappa_{\tilde{h}} = \kappa_{\alpha}(s); \tau_{\tilde{h}} = \tau_{\alpha}(s)$.

Now, consider a point $M \in \Re^3$. There exists a family of unit speed smooth space curves that pass through M. For each such curve, it is possible to construct an osculating helix that locally approximates the curve at M. Each such approximation is only valid in a small neighborhood around M (denoted $\mathcal{N}(M)$). The projection of M onto the left image plane is $m_l = P_l(M)$. Similarly, the projection of the neighborhood around M is a neighborhood around m_l, $\mathcal{N}(m_l) = P_l(\mathcal{N}(M))$, where $\mathcal{N}(m_l)$ is a connected subset of the image plane. The neighborhood of the stereo pair $(m_l, m_r) = (P_l(M), P_r(M))$ is the set of corresponding points in the space $\mathcal{N}(m_l) \times \mathcal{N}(m_r)$.

Let M be a point in $\Re^3$ and $\boldsymbol{T}_M$ the tangent vector at M. The mapping S is the projection of M and $\boldsymbol{T}_M$ onto the left and right image planes:

$$S : E^3 \times T(E^3) \mapsto E^2 \times E^2 \times T(E^2) \times T(E^2)$$

We refer to the space $E^2 \times E^2 \times T(E^2) \times T(E^2)$ as the stereo tangent space and denote it with the symbol ϑ. A point $i \in \vartheta$ can be represented as $(x_l, y_l, x_r, y_r, \theta_l, \theta_r)$ where x and y are the projection of M in the image plane and θ is the orientation of the projected tangent in $T(E^2)$. The mapping S is invertible everywhere except for the line between the optical centers and the focal planes.

Two image points, one from the left image plane and one from the right, are referred to as *corresponding points* if the inverse projection of the pair is a point on some object in the scene.

We also define the projection map:

$$\pi : E^2 \times E^2 \times T(E^2) \times T(E^2) \mapsto E^2 \times E^2$$
$$\pi(x_l, y_l, x_r, y_r, \theta_l, \theta_r) = (x_l, y_l, x_r, y_r)$$

The map $P^{-1}(\pi(i))$ maps the stereo tangent pair, i, to the corresponding point in three space. The map $S^{-1}(i)$ maps the stereo tangent pair, i, to a point in three space along with its associated tangent.

Since α at M can be approximated by an osculating helix at M, the image of the curve α in a neighborhood around $(m_l, m_r) = (P_l(M), P_r(M))$ is commensurate with the projection of the osculating helix into the left and right image planes. Furthermore, the direction of corresponding tangents in the left and right image planes is approximately equal to the direction of the projected tangents along the osculating helix. *Each pair of corresponding points and tangents is referred to as a stereo tangent pair.* If i is a stereo tangent pair, $i \in \vartheta$, associated with a point on an arbitrary curve in three space, then in a neighborhood around i, the set of stereo tangent pairs will be commensurate with the positions and tangents of the projected osculating helix. We denote the neighborhood around i as $\mathcal{N}(i)$, where $\mathcal{N}(i) = \{j \in \vartheta | \pi(j) \in \mathcal{N}(\pi(i))\}$. Formally,

STEREO COMPATIBLE: Let $i \in \vartheta$ and $j \in \mathcal{N}(i)$. $\{i, j\}$ are *stereo-compatible* under S if and only if $\exists h(s)$, a circular helix, such that $S(h(s)) \supset \{i, j\}$. We denote the compatibility of i and j under the helix h as $(i \overset{h}{\sim} j)$.

Figure 5 depicts the projection of several helices onto the left and right image planes, each of which is locally consistent (according to the co-circularity transport constraint). Equivalently, for a pair of nearby points that lie on the curve in the left image, for example m_l and n_l, the position and tangent at n_l satisfy the transport constraints at m_l. The same property holds for m_r and n_r, two nearby points on the curve in the right image plane. In addition, for each pair of neighboring stereo pairs, $m = (m_l, m_r)$ and $n = (n_l, n_r)$, the inverse projection of the corresponding points lie on the original helix and the inverse projection of the tangents at m and n have the same orientation as the tangent in three space. This is true for every pair of corresponding points that lie on the perspective projection of the helix. Since the inverse projections $P^{-1}(m)$ and $P^{-1}(n)$ lie on a common helix, the transport constraints at m are necessarily satisfied, when we transport along the image of the helix from n to m. Therefore, we can express the compatibility between a pair of corresponding points as a compatibility relationship between stereo tangent pairs. However, *one cannot arbitrarily pick a curve in the left image and a curve in the right image and expect there to be a helix in* $\Re^3$ *that projects to the given curves.* Only those helices that satisfy the compatibility relation for a given stereo tangent pair are admissible pairings of curves in left and right images. It is the difference in the projection of the helices that allows us to solve the correspondence problem. In fact, we have the UNIQUENESS LEMMA:

Let $i_1, i_2 \in \vartheta$ such that $P^{-1}(\pi(i_1)) = P^{-1}(\pi(i_2))$ and $S^{-1}(i_1) \neq S^{-1}(i_2)$. There does not exist a circular helix h and $j \in \mathcal{N}(i_1) = \mathcal{N}(i_2)$, such that $(i_1 \overset{h}{\sim} j)$ and $(i_2 \overset{h}{\sim} j)$.

Proof Suppose there exist h such that $(i_1 \overset{h}{\sim} j)$ and $(i_2 \overset{h}{\sim} j)$. From the definition of compatibility $S^{-1}(j)$ lies on the helix and the helix h passes through $M = P^{-1}(\pi(i_1)) = P^{-1}(\pi(i_2))$. We transport along the helix from $P^{-1}(\pi(j))$ to M. We denote the tangent at M as $\boldsymbol{T}_M$. Since $(i_1 \overset{h}{\sim} j)$, $i_1 \in S(h)$ and the projection of $\boldsymbol{T}_M$ is the tangent at i_1. But $(i_2 \overset{h}{\sim} j)$, so $i_2 \in S(h)$ and the projection of the

$\boldsymbol{T}_M$ is the tangent at i_2. But the tangent at $i_1 \neq$ tangent at i_2 and therefore $\boldsymbol{T}_M$ has two values.

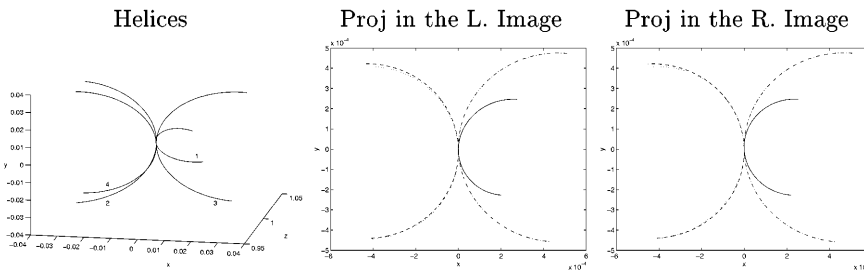

Fig. 5. Four helices in three space whose projection pass through the point $i = (0, 0, 0, 0, \pi/2, \pi/2)$. The projections in the left and right image planes are shown.

The COMPATIBILITY FIELD for $i \in \vartheta$ under S is the set of all pairs (j, h_j), where $j \in \mathcal{N}(i)$ and h_j is a unit speed circular helix, such that the relation $(i \overset{h_j}{\sim} j)$ holds for the camera model S. We denote the compatibility fields for a point i as $\mathrm{CF}_S(i)$.

The solution to the correspondence problem can now be framed in terms of the local approximation of curves. Consider a point M on a smooth curve in three space and a pair of points, not necessarily corresponding points, in the image plane $m = \{m_l, m_r\}$ (the subscripts, respectively, indicate the left and right components of the pair). Earlier, it was shown that there exists an osculating helix to the curve at M. Therefore, m_l and m_r are a match if and only if the image of the osculating helix in the left and right image planes is coincident with the locus of edge points around m_l and m_r.

Formally, solving the correspondence problem is equivalent to solving the transport problem. Consider two arbitrary curves, one in the left image plane and one in the right. Pick two tangents, (m_l, θ_l) and (m_r, θ_r) such that they lie on matching epipolar lines in the left and right image planes. The two curves can be expressed as single curve β in ϑ. Let $\beta(0) = (m_l, m_r, \theta_l, \theta_r)$ and $\beta(\delta s) = (\hat{m}_l, \hat{m}_r, \hat{\theta}_l, \hat{\theta}_r)$ be two points along β. (m_l, θ_l) and (m_r, θ_r) are corresponding tangents if there exists a helix in E^3 that satisfies the transport criteria at $\beta(0)$ and $\beta(\delta s)$. That is to say, there exists a unit speed osculating helix, $h \in \mathcal{H}$, such that $S(h(0), \boldsymbol{T}_h(0)) = \beta(0)$ and $S(h(\delta t), \boldsymbol{T}_h(\delta t)) \approx \beta(\delta s)$. These transport criteria are expressed as the position and tangent direction of the projected helix.

The compatibility field around a point $i = (m_l, m_r, \theta_l, \theta_r) \in \vartheta$ is the set of all stereo tangent pairs in a neighborhood of i for which the transport constraints can be satisfied. To draw this in a figure we separate the left components of i, (m_l, θ_l), from the right components (m_r, θ_r). The result is what appears to be a tangent field surrounding the point for which the compatibility field was constructed. Each tangent in the left compatibility field has a matching tangent

in the right field. However, there may be a disparity in both the position and orientation between the matching tangents. It is the differences in position and orientation between corresponding tangents in the left and right compatibility fields that are the key to solving the correspondence problem. This relationship is captured directly in the compatibility fields.

However, we are now at the place where the projective geometry (the camera model) interacts with the differential geometry:

OBSERVATION: Given a stereo tangent pair i, the equations used to determine the osculating helix to the curve α at $\alpha(S^{-1}(i))$, are underdetermined. For a stereo tangent pair, it is possible to invert the perspective projection matrices to obtain the position, M, and tangent of the curve in $\Re^3$. However, given only a stereo tangent pair, it is not possible to determine the direction of the normal, curvature or torsion. Therefore, there does not exist enough information to uniquely determine the helix passing through M. The family of possible helices passing through M is homeomorphic to the space $S^1 \times \mathbb{R}^+ \times \mathbb{R}$ (where S^1 denotes the unit circle). In the next section, we will provide an argument based on the principle of differential invariance to consider only a subset of the helices in $S^1 \times \mathbb{R}^+ \times \mathbb{R}$.

2.1 The Invariant Compatibility Field

The construction of a compatibility field requires a precise characterization of the camera (intrinsic/extrinsic) parameters. It would be exhausting if the compatibility fields had to be recomputed for every set of camera parameters; it would be unrealistic to assume that the camera parameters could be obtained with arbitrary precision. We seek an INVARIANT COMPATIBILITY FIELD: Consider the mappings S_1 and S_2 where each mapping is constructed using a different set of camera parameters. Let $i \in \vartheta$. For every $j \in \mathrm{CF}_{S_1}(i)$ $\exists h_j \in \mathcal{H}$ such that $S_1(h_j) \supset \{i, j\}$. The $\mathrm{CF}_{S_1}(i)$ is invariant to changes in the camera parameters if for all h that satisfy $S_1(h) \supset \{i, j\}$, $S_2^{-1} S_1(h) \in \mathcal{H}$.

To simplify the analysis, we assume that only the camera parameters θ, dx and dy are variable. Further we assume that the fixation point of the camera, (x_o, y_o, z_o) is such that $z_o > x_o$, $z_o > y_o$ and $z_o \gg$ dx. The last restriction constrains the value of θ to be close to zero. Lastly, we assume that the cameras are symmetrically verged at the fixation point.

If the compatibility field is computed for a small neighborhood around a point $i \in \vartheta$, then the affine projection equation can be used to study the invariance of the compatibility fields under changes in the camera parameters. Specifically, we determine if a helix that satisfies the compatibility criteria between two points in ϑ is preserved under changes in the camera parameters by studing the deformation of the spherical image of a circular helix under $F \triangleq A_{2*}^{-1} \circ A_{1*}$, where A_{1*} is the affine derivative map of S_1. We refer to F as the affine stereo transformation.

We can use the spherical image of a helix, σ, to define a circular cone in R^3, $\mathbf{x}_\sigma(u, v) = u\sigma(v)$. Let α be a unit speed cylindrical helix. For every parameterization β of α, the image of β' lies on a cone defined by the spherical image of α. This can be proved by defining β as $\beta(t) = \alpha(r(t))$ and consequently

$\beta'(t) = r'(t)\alpha'(r(t))$. Fix s $\in (0, L)$, where L is the arc-length of α. We can write $\beta'(t^*) = r'(t^*)\alpha'(s)$ where $t^* = r^{-1}(s)$. $\alpha'(s)$ is a point on S^2 (unit sphere) represented by the vector $\boldsymbol{v}$ so $\beta'(t^*) = r'(t^*)\boldsymbol{v}$. Thus for any reparameterization by a function r, $\beta'(t^*)$ is a point along the ray defined by $\boldsymbol{v}$ and $\boldsymbol{v}$ lies on the cone defined by the spherical image of α.

Thus an equivalence relation can be defined on the family of cylindrical helices. Each equivalence class consists of all helices that are the reparameterizations of a given unit speed cylindrical helix. The spherical image of every helix in the equivalence class lies on the cone defined by the spherical image of the unit speed helix. We require the following results:
1. The image of a circular cone under a linear mapping F is a generalized cone.
2. By the Jordan curve theorem, the curve α has a well defined interior. Consider the closed set, C, where $\alpha = \partial C$. Suppose the set C is convex. The image of this set under the linear mapping F is also convex [3]. Therefore, if $\alpha(v)$ has curvature less than zero, the curve $\hat{\alpha}(v)$ also has curvature less than zero. This implies that $\hat{\mathbf{x}}(u, v)$ is always bending away from the normal (if the cone is oriented with outward normals) and that the normal curvature is less than or equal to zero everywhere along the generalized cone. Along the rulings of the cone, the principal curvature $k_1 = 0$. The other principal curvature is k_2 is less than zero.

The above lemmas allows us to phrase the problem of invariance in terms of deformations of the circular cone:

Theorem The compatibility field is invariant if $\forall h \in \mathcal{H}$, $F(\mathbf{x}_\sigma) \in \mathcal{C}$, where $\mathcal{C}$ is the family of all circular cones, $\mathbf{x}_\sigma \in \mathcal{C}$ and σ is the spherical image of the circular helix.

A global property of a circular cone is that there exists a vector $\boldsymbol{a}$ such that the angle between $\mathbf{x}(u, v)$ and $\boldsymbol{a}$ is a constant, $\angle(\mathbf{x}, \boldsymbol{a}) = k$, $\forall u > 0, v \in \mathbb{R}$. If the cone $F(\mathbf{x})$ is a circular cone then:

$$\begin{aligned} &\angle(F(\mathbf{x}(u, v)), F(\boldsymbol{a})\ ) = k \\ &\lambda(u, v) \triangleq \frac{\partial \angle(F(\mathbf{x}(u, v)), F(\boldsymbol{a})\ )}{\partial v} = 0 \end{aligned} \tag{2}$$

The function λ is a smooth continuous function because the surface of the cone is smooth. Since λ is a function of the derivative map A_*, it is also a function of the camera parameters, e.g. θ_l and θ_r. If the camera parameters defined by A_1 are fixed, and if initially $A_2 = A_1$, the derivative of λ with respect to some camera parameter of A_2 measures the invariance of the compatibility field to changes in that camera parameter. If for example $\frac{d\lambda}{d\theta_l}$ is non-zero, then the helix is deformed by the mapping $A_2^{-1}A_1$ as θ_l varies. Therefore, λ can be used to study which unit speed circular helices are invariant to changes in the camera parameters.

Consider the set of all unit speed circular helices, $\mathcal{H}_0$, that lie on a circular cylinder whose axis goes through the origin. Let the direction of the axis represent the direction of the helix. The direction of every helix in $\mathcal{H}_0$ can be represented by a point on S^2. We can change the direction of each helix with a

rotation matrix R, which has two degrees of freedom. Consider the point $(0, 0, 1)$ on S^2. Any other point on S^2 can be defined by a rotation around the x-axis followed by a rotation around the y-axis. We define the mapping $R : S^2 \mapsto S^2$ as such a function. The mapping R is a smooth continuous function [3]. The mapping R changes the direction of a helix and rotates the spherical image of the helix accordingly. If σ_1 is the spherical image of unit speed helix, $h_1 \in \mathcal{H}_0$, and we define $h_2 = R(h_1)$, then by the linearity of differentiation the spherical image of h_2 is $\sigma_2 = R(\sigma_1)$.

Equation 2 can be used to study the invariance of the helix h_1 by considering its spherical image σ_1. Similarly, it is possible to study the invariance of h_2 using the spherical image $\sigma_2 = R(\sigma_1)$. Therefore, it is possible to study the invariance of an entire family of helices, each with a different direction by considering different rotation matrices. There is a natural symmetry on S^2 (unit sphere) based on the fact that the deformation associated with a point on S^2 and its conjugate are equal. Figure 6 shows an example of the surface that is generated by considering the deformation for a family of helices, a subset of $\mathcal{H}_0$, that have the same value of curvature and torsion but whose directions vary. The locus of points along the bottom of the valley represent the helix whose axis is defined by the vector $(0, 0, 1)$. We will denote this helix as the z-helix. The figure demonstrates that the z-helix is most stable to small changes in the camera parameter θ_l. A similar observation is made when other camera parameters are varied.

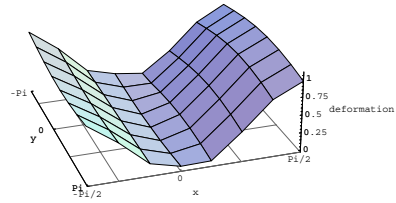

Fig. 6. The degree of deformation, normalized between 0 and 1, for helices pointing in different directions. The direction of each helix is represented by a rotation around the x and y axis of the standard Euclidean frame. The variables x and y represent a rotation about the x and y axis respectively. The helix with 0 deg of rotation around the x-axis is referred to as the z-helix. The value of curvature and torsion for each helix in the family is $\frac{1}{\sqrt{2}}$ and $\frac{1}{\sqrt{2}}$. See text for details.

Figure 7:Top shows three helices which are oriented along three different and mutually perpendicular axes. The degree of invariance to changes in θ_l for each of the helix directions are shown at the bottom of the figure. Of the three helices, the z-helix is the most stable to changes in the camera parameters for most ratios of $\frac{\tau}{\kappa}$. A similar observation is made for small changes in θ_r, dx and dy.

It was stated earlier that any solution to the stereo correspondence problem must satisfy the principle of differential invariance. In figure 7 the magnitude of

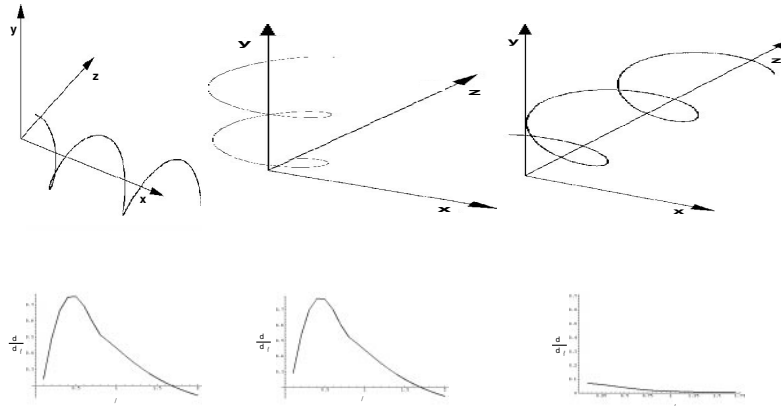

Fig. 7. Top: Shown are three helices in E^3, denoted x-helix, y-helix, z-helix from left to right respectively. Bottom: The degree of deformation of each helix type as the left camera angle is slightly perturbed. The maximum amount of deformation of each helix type is shown as a function of the ratio τ/κ. The line of sight is along the z-axis. The z-helix, the rightmost helix, is significantly more stable under camera perturbations than the other two.

the maximum deformation of the x- and y-helix are significantly larger compared to the z-helix. Therefore, of the three helices, the z-helix comes closest to satisfying the principle of differential invariance.

The orientation disparity encoded by a stereo tangent pair can now be interpreted in terms of the physical behavior of a z-helix in three space. If the amount of torsion is equal to zero, the helix is confined to the plane of fixation and the orientation disparity is necessarily zero [9]. As the torsion increases, the z-helix begins to twist out of the plane of fixation. The tangent at the point where the helix crosses the plane of fixation is no longer parallel to the plane. This difference induces an orientation disparity between the projection of the tangent in the left eye and the tangent in the right eye.

Two technical details remain, that are beyond the space limitations in this paper. First, the choice of the z-helix implies that it can be used as the local approximation to an arbitrary space curve. Unfortunately, this requires the Frenet frame of the z-helix to be oriented in the same direction as the curve's frame. However, it can be shown that, for appropriately selected values of curvature and torsion, the z-helix is a valid approximation to a curve under the limits induced by the quantization of position, orientation and curvature in the image planes. Second, the sampling of the stereo tangent space may introduce complications because the inverse projection of a quantized stereo pair may not be defined. However, it can be shown that, for suitable couplings between positon and orientation quantization, a suitable pseudoinverse can be defined (Fig. 8).

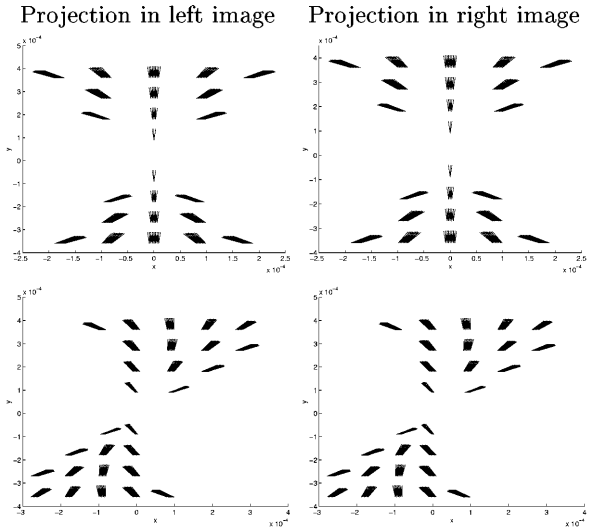

Fig. 8. Top: The positive discrete compatibility field for the point $\hat{\imath} = (0, 0, 0, 0, \pi/2, \pi/2)$ projected in the left and right image planes. Bottom: The positive discrete compatibility field for the point $\hat{\imath} = (0, 0, 0, 0, \pi/3, \pi/3)$ projected in the left and right image planes.

2.2 The Stereo Relaxation Labeling Process

We can now use the discrete compatibility fields to select pairs of points in the stereo tangent space that are most consistent with one another. For motivation it helps to recall the model for monocular edge consistency via relaxation labeling with co-circular compatibilities. Abstractly, let a set of nodes (image positions) be given, and a set of labels (edge orientations) be defined for each position. Further, let the labels be ordered at each node by a probability measure $p_i(\lambda)$, indicating the probability that label λ is correct for node i, with $\sum_\lambda p_i(\lambda) = 1 \forall i$. Labels are "selected" at each position by an iterative gradient ascent that extremizes the functional $A(p) = \sum p_i(\lambda) r_{i,j}(\lambda, \lambda') p_j(\lambda')$ in parallel for all nodes i and all labels λ. The compatibilities $r_{i,j}(\lambda, \lambda')$ derive from co-circularity[15]. For full description see [8]. In effect, the relaxation network selects those edges that, if transported along the osculating circle locally, would most match their neighbors.

A specialization of the above to the two-label relaxation labeling process was developed to allow multiple labels at a position[15], and it is this specialization that will be used to disambiguate the stereo tangent pairs. Let $\mathcal{I} = \{1, \ldots, n\}$ be a set of nodes. Each node is assigned two labels $\lambda \in \{\text{TRUE,FALSE}\}$ where $p_i(\lambda)$ represents the confidence in the label at node i and $p_i(\text{TRUE}) + p_i(\text{FALSE}) = 1$. Since it is sufficient to know the value of $p_i(\text{TRUE})$ to determine $p_i(\text{FALSE})$, only values of $p_i(\text{TRUE})$ need be updated. For notational convenience, let $p_i =$

p_i(TRUE). The update rule is:

$$\begin{aligned} p_i^{t+1} &= [p_i^t + q_i]_0^1 \\ &= [p_i^t + \sum_j r_{ij} p_j^t]_0^1 \end{aligned} \tag{3}$$

where

$$[x]_0^1 = \begin{cases} 1, & \text{if } x > 1 \\ x, & \text{if } 0 \leq x \geq 1 \\ 0, & \text{if } x < 0 \end{cases}$$

We denote q_i as the support at p_i and $R = \{r_{ij}\}$ as the compatibility matrix.

We use the two-label relaxation process to infer the correspondences between tangents in the left and right tangent fields. Let each kernel point, $\hat{\imath} \in \hat{\vartheta}$, be a node in the relaxation labeling network, where the $\hat{}$ denotes discretization. Thus, each node in the network encodes absolute horizontal disparity, vertical disparity, and orientation disparity. Intuitively, the value of $p_{\hat{\imath}}$(TRUE) represents the confidence that the stereo tangent pair $\hat{\imath}$, is a correct correspondence. The initial labeling assignment is the set of all possible correspondences in the image plane. The value of the compatibility matrix element $r_{\hat{\imath}\hat{\jmath}}$ is the discrete compatibility between the pair $\hat{\imath}$, $\hat{\jmath}$. As a consequence, the set of all discrete compatibility fields is commensurate with the compatibility matrix R.

3 Results

The first step in our stereo process is to obtain a discrete tangent map representation of the stereo pair. Figure 9 is the reconstruction of the scene from the synthetic stereo pair using our approach to correspondence matching (without curvature information). In those few places where the differential geometric constraints are inadequate, we select those matches that lie closest to the fixation plane. The most significant observation with respect to the reconstruction is that the change in the ordering of primitives, in this case discrete tangent elements, does not effect the matching algorithm. Two curves, a y-helix and an arc, are both successfully segmented and rendered at the correct depth. Another significant difference between the TINA reconstruction and ours is that our curves are both isolated in space. This difference is due in part to the failure of the PMF matching algorithm near junctions in the image. We handle such structures in the discrete tangent map by representing multiple tangents at a given (image) point. At image junctions, the tangent map encodes two separate tangent directions, one for the arc and the other for the helix. The multiple tangents at the junction in the left image participate in the matching process with the appropriate tangents at the junction in the right image. This ensures that the parts of the helix in the scene that correspond with the junction areas in the image have valid matches in the left and right image planes. The same fact also applies for the arc.

The errors in matching induced by the lack of ordering are further exaggerated when the helix is farther out in front of the circular arc. This is because the image extent increases over which the ordering constraint is invalid. Fig. 2 contains the same curves as the previous stereo pair but exaggerates the parts of the image where the ordering constraint breaks down. Predictably, the number of pixels that are falsely matched using the PMF algorithm increases (Fig. 3). The reconstruction of the scene as computed by TINA is shown in Fig. 3.

In contrast to the TINA reconstructions, our stereo algorithm returns a more veridical reconstruction of the original scene (see figure 9). As in the previous stereo pair, the only constraint used in the computation was a bias for matches close to the fixation plane. The stereo reconstruction results in two curves, an arc and a y-helix that are separated in depth.

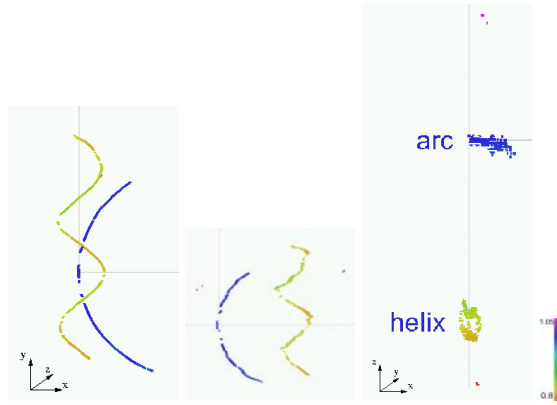

Fig. 9. The reconstruction of the stereo scene from the stereo pair in Fig. 2 is shown from three separate projections. The color of each space tangent represents its depth (red is closest to the camera and blue the farthest away). Note that the approximate geometry of the curves in three space is reconstructed and that the curves are localized at distinct depths. The result is obtained after 5 iterations of relaxation labelling and $\sigma_c = 1.75$.

Fig. 10. A stereo pair of Asiatic lilies. The cameras had a 9cm baseline. The stereo pair can be fused using uncrossed fusing.

Fig. 11. The discrete tangent map of the lilies in Fig 10. The tangent map was created using logical/linear operators [10] and a relaxation labelling process [15] followed by a deblurring function (after [4]). The quantization of space and orientation has been improved by six times in both dimensions.

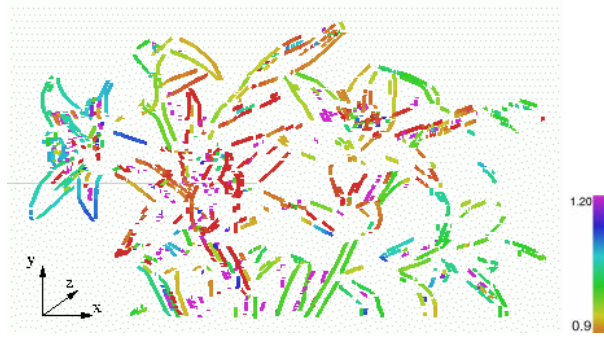

Fig. 12. The depth map associated with the stereo lily pair. The color bar at the right indicates the depth associated with the colors in the image. Each point in the image is a tangent in three space that is geometrically consistent with its neighbors.

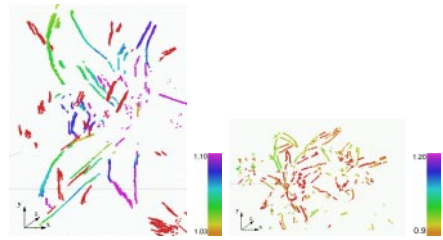

Fig. 13. Left: A detail of a flower from the stereo reconstruction. The colored depth map is shown to the right of the image. The red curve segments are false matches that are spatially located in front of the flower. The graduated coloring along the petals indicates a smooth transition along the petals. Right: A cropping plane 1.01m away from the cyclopean center is inserted to show the details of objects in the foreground.

The gaps in the stereo reconstructions are due mainly to the inability match horizontal tangents. Image tangents whose orientations are less than 20deg from the horizontal are not used in the stereo reconstruction. There are some false matches as shown in the left panel of the above figures.

We have shown that a key assumption used in algorithms like PMF is the ordering constraint. In the synthetic images we showed that if the ordering constraint is violated then the PMF algorithm performs poorly. In contrast, our stereo algorithm is relatively unaffected by the lack of ordering in parts of the image. The stereo pair of twigs in Fig. 1 shows a real world situation where the ordering constraint breaks down but where our algorithm succeeds (Fig. 14). While there are false matches associated with the twigs farthest in the background, they are few in number and are often short in length. They occur in this implementation because preference is given to matches closest to the fixation plane which is in front of the receding twigs.

Fig. 14. Left: The stereo reconstruction of the twigs scene. There are some mismatches near the base of the twigs and along the branch receding in depth. The reconstruction is not affect by those areas where there is failure of the ordering constraint. Middle: A cropping plane 1.03m from the cyclopean point was placed to highlight the twigs in the foreground. Right: A birdseye view of the reconstructed scene. The twigs are arranged in roughly three layers as can be seen by fusing the stereo pair.

In summary, we have shown how the differential geometry of curves in monocular images can be made consistent with (the Frenet estimate of) the spatial curve from which they project to define a relaxation labeling algorithm for stereo correspondence. In effect we have generalized co-circularity for planar curves to co-helicity for spatial curves. Invariance to camera-model uncertainties dictates the z-helix from among a family of possible space helices, and a unique feature of this correspondence is the natural way that orientation disparity combines with positional disparity. Examples illustrate how our system functions even in situations when ordering (and other heuristic) constraints break down. The evolutionary pressures on early tree-dwelling mammals, in particular early primates, suggests that uncertainties in reaching for a branch while swinging across space could not be tolerated. Given the mapping of our stereo tangent space onto orientation hypercolumns, perhaps our algorithm also implies a mechanism for stereo correspondence in visual cortex.

References

1. H. H. Baker and T. O. Binford. Depth from edge and intensity based stereo. In *Proceeding of the 7^{th} Joint Conference of Artificial Intelligence*, pages 631–636, 1981.
2. S.T. Barnard and W.B. Thompson. Disparity analysis of images. *PAMI*, PAMI-2(4):333–340, July 1980.
3. E.D. Bloch. *A First Course in Geometric Topology and Differential Geometry.* Birkhauser, Boston, MA., 1997.
4. C. David and S.W. Zucker. Potentials, valleys and dynamic global coverings. *International Journal of Computer Vision*, 5(3):219–238, 1990.
5. Olivier Faugeras. Three dimensional computer vision. *MIT Press*, 1993.
6. J.P. Frisby and S.B. Pollard. Computational issues in solving the stereo correspondance problem. In M.S. Landy and J.A. Movshon, editors, *Computational Models of Visual Processing*, chapter 21, pages 331–357. MIT Press, Cambridge, MA, 1991.
7. W.E. Grimson. A computer implementation of a theory of human stereo vision. In *Philisophical Transactions of the Royal Society of London*, volume B292, pages 217–253. 1981.
8. Robert Hummel and Steven Zucker. On the foundations of relaxation labelling processes. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 5(3):267–287, 1983.
9. B.J. Rogers I.P. Howard. *Binocular Vision and Stereopsis.* Oxford Psychology Series No. 29. Oxford, 1995.
10. Lee Iverson and Steven Zucker. Logical/linear operators for image curves. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17(10):982–996, October 1995.
11. Jesse Jin, W.K. Yeap, and B.G. Cox. A stereo model using log and gabor filtering. *Spatial Vision*, 10(1):3–13, 1996.
12. David Jones and Jitendra Malik. Determining three-dimensional shape from orientation and spatial frequency disparities. *ECCV 92*, pages 661–669, 1992.
13. David Marr and T. Poggio. A theory of human stereo vision. *Proceedings of the Royal Society of London*, B 204:301–328, 1979.
14. G. Medioni and R. Nevatia. Segment-based stereo matching. *Computer Vision, Graphics and Image Processing*, 31:2–18, July 1985.
15. Pierre Parent and Steven Zucker. Trace inference, curve consistency and curve detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(8):823–839, August 1989.
16. Doron Sherman and Shmuel Peleg. Stereo by incremental matching of contours. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12(11):1102–1106, November 1990.