

# A Differential Geometrical Model for Contour-Based Stereo Correspondence

Gang Li and Steven W. Zucker  
Department of Computer Science  
Yale University, New Haven, CT 06520  
{gang.li, steven.zucker}@yale.edu

## Abstract

*Standard approaches to the computation of stereo correspondence have difficulty when scene structure does not lie in or near the frontal-parallel plane, in part because an orientation disparity as well as a positional disparity is introduced. We propose a correspondence algorithm based on differential and projective geometry, and inspired by neurobiology, that takes explicit advantage of both disparities. Based on curves, the algorithm relates the (2D Frenet) differential structures (position, tangent, and curvature) in the left and right images with the Frenet geometry of the (3D) space curve. A compatibility function is defined via transport of the Frenet descriptors, and they are matched by relaxing this compatibility function on overlapping neighborhoods along the curve. False correspondences are concurrently eliminated by a model of "near" and "far" neurons derived from neurobiology. Examples of our algorithm and standard approaches are compared.*

## 1. Introduction

A glance at Fig. 1 reveals many of the problems facing stereo correspondence algorithms [18]: structure is not fronto-parallel, as in many man-made scenes; ordering constraints [16] are violated accordingly; uniqueness constraints [12, 16, 23] are inconsistent with self-occlusion, branching and discontinuities are inconsistent with smoothness constraints [12, 13]; and feature density is not uniform. Nevertheless, there is a structural relationship between features in the (left, right) image pair and the 3D scene, and our goal in this paper is to develop this relationship into a new stereo-correspondence algorithm.

There are two sources of motivation for our algorithm. The first is biological: tree-dwelling primates are readily able to solve for correspondence in images such as Fig. 1 using binocularly-selective neurons that are also orientation selective. This suggests an approach based on differential geometry, with orientation identified with the tangent to spatial structure. Our second motivation develops this mathematical interpretation, and we use projective differential geometry to derive the underlying structural relationship

between image features and 3D spatial structure. But even with the epipolar constraint there still exist ambiguities: different pairings of image features give rise to different 3D structures, thereby creating multiple ghost matches as well as correct ones. So finally we return to the biological motivation, and introduce an extended model for binocular neurons that implements their "near" and "far" tuning properties. The result is a relaxation network that eliminates ghost matches and explains mechanistically the disparity gradient limit [16].

Our geometric calculations follow the literature on curve matching. Cipolla and Zisserman [3] determined image tangent and geodesic curvature from a single view of a space curve under perspective projection. Faugeras and Robert [17, 6] used tangent and curvature constraints in a trinocular system, and predicted the curvature at a point in the third image based on measurements in the first two. Assuming knowledge of the fundamental matrix, Schmid and Zisserman [19] described how to compute the normal at one 3D space curve point, thus determining the osculating plane at that point from corresponding tangents and curvatures at two perspective image points. We extend this previous work by showing how to compute both the normal and the curvature at one 3D space curve point from the perspective projections of the space curve in two views.

These normals and curvatures provide the basis for our algorithmic framework, which follows the relaxation or belief propagation model. Compatibility fields are defined from a local (Frenet) approximation to a space curve, and these are used to determine how compatible two pairs of candidate matches are. Shan and Zhang [20] used combined binary measurements for line segments and curves as the compatibility function. [1] approximated the curve heuristically with a z-helix, and they used a disparity bias (to the plane of fixation) to eliminate false matches. We individually compute the space curve local approximation for each candidate match pair using curvature, to eliminate the need for heuristics. A minimum torsion approximation guides the matching, and we integrate the near and far neuron mechanism to address the remaining ambiguity. To our knowledge this is the first time such mechanisms have been used in stereo matching, and they eliminate the need for a

disparity bias.

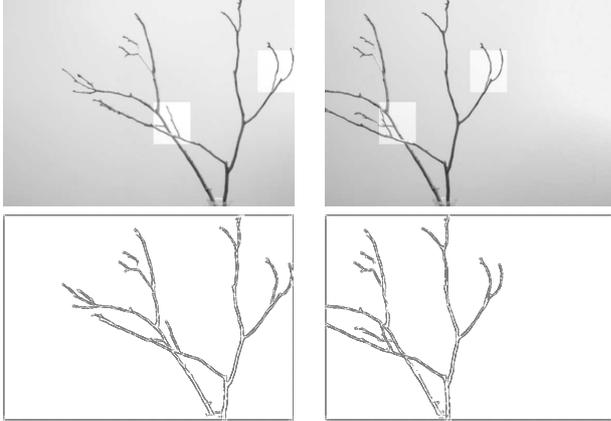


Figure 1: (top) Natural scenes are dense in physical structure with depth discontinuities, occlusions, and junctions, thereby posing a problem to traditional stereo algorithms. For the region around the center of the left image, when compared to its corresponding part in the right image, note that the ordering constraint fails because of the depth discontinuity. Partial occlusion at the junction would cause a problem for the uniqueness constraint. Our algorithm is designed to function on image pairs such as this. (bottom) The (left,right) tangent (edge) map pair for the (top) images. Such tangent maps form the basis for our correspondence algorithm, and suggest a curve-based approach.

## 2. Geometry of Space Curves

Let  $\alpha(s)$  be a regular curve parametrized by arc length in  $R^3$ . If the Frenet (tangent, normal, binormal) frame  $\{ \mathbf{T}_0, \mathbf{N}_0, \mathbf{B}_0 \}$ , the curvature  $\kappa_0$ , and torsion  $\tau_0$  are known at  $\alpha(0)$ , we can obtain a local Frenet approximation of the curve by taking the third-order Taylor expansion of  $\alpha$  at  $s = 0$  and keeping only the dominant terms:

$$\hat{\alpha}(s) = \alpha(0) + s\mathbf{T}_0 + \frac{s^2}{2}\kappa_0\mathbf{N}_0 + \frac{s^3}{6}\kappa_0\tau_0\mathbf{B}_0 \quad (1)$$

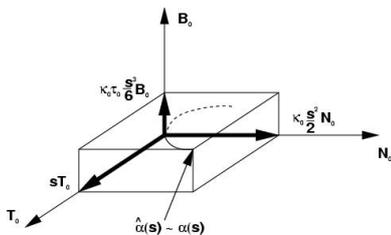


Figure 2: Frenet approximation of 3D curve around  $\alpha(0)$ .

The projection operator  $\pi$  maps this local frame to the left and right image planes:

$$\pi : R^3 \times T(R^3) \times N(R^3) \times B(R^3) \times \kappa(R) \times \tau(R) \mapsto R^2 \times T(R^2) \times \kappa(R) \times R^2 \times T(R^2) \times \kappa(R)$$

That is, the Frenet trihedron in  $R^3$  projects to two Frenet (tangent, normal)-dihedra, one in the left image and one in the right image, each augmented by curvature. We refer to the space  $R^2 \times R^2 \times T(R^2) \times T(R^2) \times \kappa(R) \times \kappa(R)$  as the *stereo tangent space*; a point in this space is  $(x_l, y_l, x_r, y_r, \theta_l, \theta_r, \kappa_l, \kappa_r)$ , where  $x_i$  and  $y_i$  are the image projection coordinates of  $X \in R^3$ ,  $\theta_i$  the orientation of projected tangent in the image planes, and  $\kappa_i$  the image curvatures, with  $i = \{ l, r \}$  representing left and right images.

### 2.1. 3D Space Curve Structure From Two Views

In this section our technical goal is to determine the inverse mapping  $\pi^{-1}$ , the space curve structure around the 3D point  $X$  from its projection in the stereo tangent space. Unfortunately, as we will show,  $\pi^{-1}$  is not one-to-one: given a node in the stereo tangent space  $i = (x_l, y_l, x_r, y_r, \theta_l, \theta_r, \kappa_l, \kappa_r)$ , we can only determine the position  $\mathbf{X}$ , the Frenet frame  $\{ \mathbf{T}, \mathbf{N}, \mathbf{B} \}$ , and the curvature  $\kappa$  at the space curve point  $X$ . The torsion  $\tau$  can not be determined. To prove this we assume the following image measurements are given: position, tangent, curvature, i.e.  $(x, y, \theta, \kappa)$  at every image curve point in both images. We further suppose (for space reasons) that the 3D position  $\mathbf{X}$  and tangent  $\mathbf{T}$  are computed by standard methods [5] or [7]. We now describe how to compute the normal and curvature at a 3D space curve point from two views.

We begin with a standard construction [2]. Let  $\alpha$  be a smooth space curve with non-vanishing curvature  $\kappa$ . Denote its position by the vector  $\mathbf{r}(s)$  in the world coordinate system, and its spherical projection to the unit radius image sphere with center  $\mathbf{c}$  by a vector  $\mathbf{u}$  in the camera coordinate system (see Fig. 3). Assume the world and the camera coordinate systems have the same orientation, that is, there is only translation, no rotations between them. Then the space curve can be described as:  $\mathbf{r}(s) = \mathbf{c} + \lambda(s)\mathbf{u}(s)$ , where  $\mathbf{c}$  is the vector pointing to the camera center in the world coordinate system,  $\lambda$  is the Euclidean distance of  $\mathbf{r}(s)$  to the camera coordinate system origin, and  $\mathbf{u}$  is the vector pointing to its spherical projection in the camera coordinate system.

Previous work [3, 2] shows that the relationship of image geodesic curvature and the 3D space curve geometry under perspective projection is:

$$\kappa^g = \frac{\lambda\kappa(\mathbf{u} \times \mathbf{T}) \cdot \mathbf{N}}{(1 - (\mathbf{u} \cdot \mathbf{T})^2)^{\frac{3}{2}}} \quad (2)$$

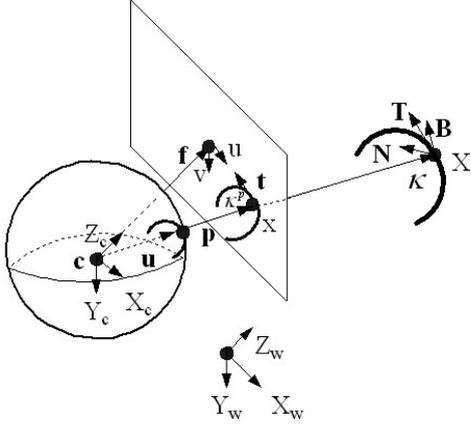


Figure 3: 3D space curve, its spherical projection to unit sphere, and its perspective projection to image plane.

where  $\kappa^g$  is the image geodesic curvature at the image spherical projection  $\mathbf{u}$  of the space point being studied,  $\mathbf{T}$  and  $\mathbf{N}$  are the space tangent and normal at that point,  $\kappa$  is the curvature at the space point  $\mathbf{r}(s)$ . We now calculate the relationship between 3D space curve properties and their (measurable) perspective projection image plane properties.

Consider the 2D image plane curve formed by  $\mathbf{r}(s)$  through perspective projection, with  $\mathbf{p}$  the vector pointing to the projection in the image plane ( $\mathbf{u} = \mathbf{p}/\|\mathbf{p}\|$ ), and  $\mathbf{f}$  the vector pointing to the image center, both in the camera coordinate system.  $\mathbf{t}$  and  $\mathbf{n}$  are the tangent and normal vectors in the image coordinate system, respectively. Since we choose the camera and image coordinate systems such that they are aligned, we can add a third component zero to the two dimensional vector  $\mathbf{t}$  and  $\mathbf{n}$  to express the tangent and normal vectors in the camera coordinate system, which is 3D. Note that this construction will form the same spherical projection as  $\mathbf{r}(s)$ . Furthermore, although for a plane curve the curvature is defined as the signed curvature  $\kappa^p$ , and the normal is defined by rotating tangent  $\pi/2$  counter-clockwise, the quantity  $\kappa^p \mathbf{n}$  is still the same if we study this curve as a 3D curve, and that the binormal at any point in this curve is  $\mathbf{b} = \mathbf{t} \times \mathbf{n} = -\mathbf{f}/\|\mathbf{f}\|$ . The relationship between the perspective projection image curvature and the geodesic curvature is thus given by:

$$\begin{aligned} \kappa^g &= \frac{\|\mathbf{p}\| \kappa^p (\mathbf{u} \times \mathbf{t}) \cdot \mathbf{n}}{(1 - (\mathbf{u} \cdot \mathbf{t})^2)^{\frac{3}{2}}} = \frac{\|\mathbf{p}\| \kappa^p \mathbf{u} \cdot (\mathbf{t} \times \mathbf{n})}{(1 - (\mathbf{u} \cdot \mathbf{t})^2)^{\frac{3}{2}}} \\ &= \frac{-\|\mathbf{p}\| \kappa^p \mathbf{u} \cdot \frac{\mathbf{f}}{\|\mathbf{f}\|}}{(1 - (\mathbf{u} \cdot \mathbf{t})^2)^{\frac{3}{2}}} = \frac{-\frac{1}{\|\mathbf{f}\|} \kappa^p \mathbf{p} \cdot \mathbf{f}}{(1 - (\frac{\mathbf{p}}{\|\mathbf{p}\|} \cdot \mathbf{t})^2)^{\frac{3}{2}}} \\ &= \frac{-\|\mathbf{f}\| \kappa^p}{(1 - (\frac{\mathbf{p}}{\|\mathbf{p}\|} \cdot \mathbf{t})^2)^{\frac{3}{2}}} \end{aligned} \quad (3)$$

Using the geodesic curvature as the bridge, we can now

connect the above two equations to formulate the relationship of the 3D space curve curvature and its perspective projection image curvature:

$$\frac{\frac{\lambda}{\|\mathbf{p}\|} (\mathbf{p} \times \mathbf{T}) \cdot \mathbf{N}}{(1 - (\frac{\mathbf{p}}{\|\mathbf{p}\|} \cdot \mathbf{T})^2)^{\frac{3}{2}}} \kappa = -\frac{\|\mathbf{f}\|}{(1 - (\frac{\mathbf{p}}{\|\mathbf{p}\|} \cdot \mathbf{t})^2)^{\frac{3}{2}}} \kappa^p \quad (4)$$

**Remark** By assuming the osculating plane (spanned by  $\mathbf{T}$ ,  $\mathbf{N}$ ) at the space curve point coincides with the  $XY$  plane of the world coordinate system (as in [22]), and using similar techniques as [19], we could also derive the relationship of the 3D space curve and image curvatures in a projective geometry framework. The world coordinate system and the camera coordinate system could have different orientations, with a homography used to relate the image plane and the osculating plane. We omit the formulas due to space limitation.

Now we can compute the normal  $\mathbf{N}$  and curvature  $\kappa$  at a 3D space curve point from image measurements  $(x_l, y_l, x_r, y_r, \theta_l, \theta_r, \kappa_l, \kappa_r)$  of one corresponding pair, where  $\mathbf{f}_i$  is obtained from calibration,  $\mathbf{p}_i$ ,  $\mathbf{t}_i$  can be easily computed from the above image measurements, with  $i = \{l, r\}$ .

**Proposition:** *Given two perspective views of a 3D space curve with full calibration, the normal  $\mathbf{N}$  and curvature  $\kappa$  at a space curve point  $X$  are uniquely determined from the positions, tangents, and curvatures of its projections in two images. Thus the Frenet frame  $\{\mathbf{T}, \mathbf{N}, \mathbf{B}\}$  and curvature  $\kappa$  at the space curve point  $X$  can be uniquely determined.*

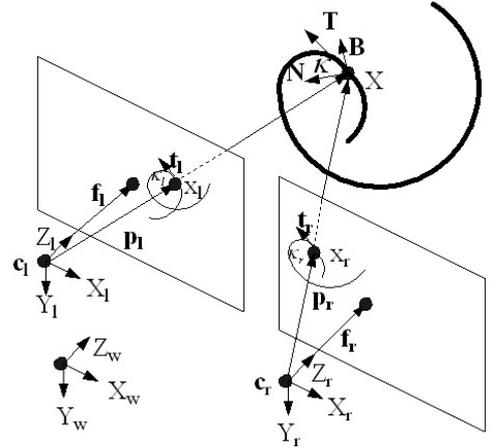


Figure 4: The geometry for Proposition. Given  $(x_l, y_l, x_r, y_r, \theta_l, \theta_r, \kappa_l, \kappa_r)$ , we can compute normal  $\mathbf{N}$  and curvature  $\kappa$  at the 3D space curve point. Thus we can determine the position, the Frenet frame  $\{\mathbf{T}, \mathbf{N}, \mathbf{B}\}$ , and the curvature  $\kappa$ . But the torsion  $\tau$  can not be determined.

Rewriting eq. (4) to keep the unknown normal  $\mathbf{N}$  and curvature  $\kappa$  on the left side, and using it for the left and right

images, the relationship of 3D position, tangent, normal, and curvature at a space curve point and its image positions, tangents, and curvatures under perspective projection is:

$$\begin{aligned} (\mathbf{p}_l \times \mathbf{T}) \cdot \mathbf{N} \kappa &= -\frac{\|\mathbf{f}_l\| \|\mathbf{p}_l\| (1 - (\frac{\mathbf{p}_l}{\|\mathbf{p}_l\|} \cdot \mathbf{T})^2)^{\frac{3}{2}}}{\lambda_l (1 - (\frac{\mathbf{p}_l}{\|\mathbf{p}_l\|} \cdot \mathbf{t}_l)^2)^{\frac{3}{2}}} \kappa_l^p \\ (\mathbf{p}_r \times \mathbf{T}) \cdot \mathbf{N} \kappa &= -\frac{\|\mathbf{f}_r\| \|\mathbf{p}_r\| (1 - (\frac{\mathbf{p}_r}{\|\mathbf{p}_r\|} \cdot \mathbf{T})^2)^{\frac{3}{2}}}{\lambda_r (1 - (\frac{\mathbf{p}_r}{\|\mathbf{p}_r\|} \cdot \mathbf{t}_r)^2)^{\frac{3}{2}}} \kappa_r^p \quad (5) \\ \mathbf{T} \cdot \mathbf{N} \kappa &= 0 \end{aligned}$$

where subscript  $l$  and  $r$  represent measurements from left and right images, respectively. The last equation specifies  $\mathbf{N}$  has to be orthogonal to tangent  $\mathbf{T}$ . This system can be solved for  $\mathbf{N} \kappa$ , from which we can compute the curvature  $\kappa$  and normal  $\mathbf{N}$ .

Figure 4 illustrates that from two views we can compute the 3D space curve normal, thus determine the Frenet frame  $\{\mathbf{T}, \mathbf{N}, \mathbf{B}\}$  and curvature  $\kappa$  at a 3D space curve point, but not the torsion  $\tau$ .

### 3. Stereo Correspondence Algorithm

In the previous section, we described how to compute the space curve local approximation given image measurements in two perspective views. Now we describe how this geometric relationship can be used to solve the stereo correspondence problem. The central idea is as follows.

At each point  $\alpha(s_0)$  along the space curve, a Frenet approximation can be constructed to characterize the local behaviour of the curve *around* that point based solely on the information *at* that point. Since measured information is also available for nearby points, this can be transported along the *approximation* to the neighborhood of  $\alpha(s_0)$  and then compared with the frame at  $\alpha(s_0)$ . That transported information which is "close to" the measured information is supported, and that which is "far" is not supported. Both position and orientation information are used in determining these distances.

The situation is analogous to co-circularity [14] for image curves. A unit speed planar curve  $\beta(s)$ , with positive curvature, has a unique osculating circle that approximates  $\beta$  around  $s$ . This approximation is used to define co-circularity constraints so that nearby (estimated) edge elements can be transported along the curve locally using its osculating circle, and a relaxation labeling network extremizes a global functional of how well the transported edges match. Compatibility functions encode these transport operations.

Three distances are used in building compatibility functions for co-circularity: the transport distance along the osculating circle, the distance from this point to the edge, and

the angular difference between the measured and the transported edges. We now seek to do this for space curves, but the situation is a little more complicated.

#### 3.1. The Minimal Torsion Constraint

To infer the 3D space curve, we first get its local approximation at each point, and then transport (estimated) tangents to nearby positions. A relaxation labeling network will infer the space curve geometry.

The heart of this process is cartooned in Fig. 5, where two space tangents along the 3D Frenet approximation are shown. Observe that each of the space tangents projects to a pair of image tangents, so the nodes  $i$  and  $j$  in the relaxation network consist in *pairs* of tangents, one in the left image and one in the right, and compatibilities  $r_{ij}$  are defined over these pairs. Observe further that, while the expected positional disparity between these tangents is introduced, there are also orientation disparities; such higher-order disparities contribute to the matching process and also have biological counterparts [8].

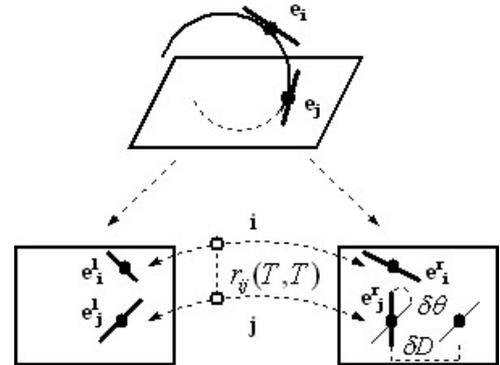


Figure 5: Cartoon of the stereo relaxation process. (top) shows a pair of space tangents associated with the Frenet approximation around the point  $e_j$ . Each of these tangents projects to a (left,right) image tangent pair; compatibility between the space tangents thus corresponds to compatibility over (left,right) image tangent pairs. The projected tangents are shown as thick lines. One left image tangent is redrawn in the right image (as thin lines) to illustrate positional disparity ( $\delta D$ ) and orientation disparity ( $\delta \theta$ ).

A stereo tangent pair, or point in the stereo tangent space, consists in the 8-tuple:  $(x_l, y_l, x_r, y_r, \theta_l, \theta_r, \kappa_l, \kappa_r)$ . The set of all stereo tangent pairs  $j$  in the neighborhood of  $i$ , such that  $i$  and  $j$  are compatible, is called the *compatibility field* around  $i$ .

To build the compatibility fields for stereo, recall that torsion remains a free parameter. Varying this term yields a family of local approximations; Fig. 6.

The Frenet frame does specify the osculating plane in  $R^3$ , however, so the transport distance that we use to define

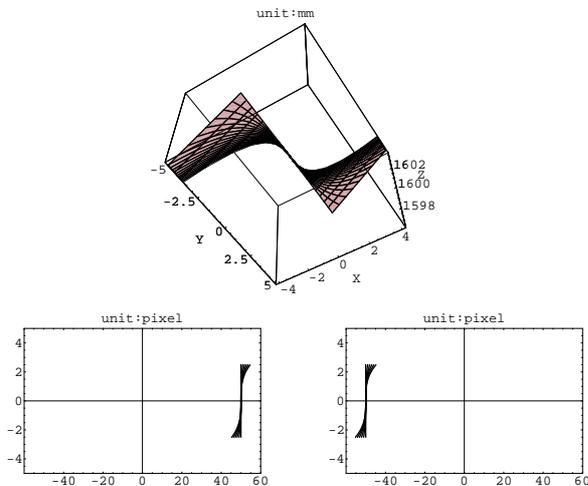


Figure 6: (TOP) Family of local approximations in 3D obtained by varying torsion. (BOTTOM) Discrete Compatibility Field projected to the left and right image planes. Both for  $i = (50, 0, -50, 0, \pi/2, \pi/2, 0.7140, 0.6188)$ , focal length 800 pixels, baseline 200mm.

the compatibility field is piecewise co-circularity in the osculating plane at each point (see Fig. 7). Since the change of the pose of the osculating plane in  $R^3$  is related to the torsion, the result amounts to a minimal torsion constraint. We conjecture that  $\tau$  varies smoothly along the curve.

### 3.2. Stereo Relaxation Labeling

We now specify the stereo algorithm in more detail. The task is to select pairs of *nodes* in the stereo tangent space that are most consistent with one another by stereo relaxation labeling using the discrete compatibility fields just sketched. (They will be refined below.)

Relaxation labeling processes assign labels to nodes in a graph based on the parallel use of local constraints. Suppose a set of nodes are given, and a set of labels are defined for each node.  $p_i(\lambda)$  is the probability that label  $\lambda$  is correct for node  $i$ , with  $\sum_{\lambda} p_i(\lambda) = 1$  for every node  $i$ . Labels are selected at each node by an iterative gradient ascent that extremizes the functional  $A(p) = \sum p_i(\lambda) r_{ij}(\lambda, \lambda') p_j(\lambda')$  in parallel for all nodes  $i$  and labels  $\lambda$  in the network. The compatibilities  $r_{ij}(\lambda, \lambda')$  specify the local constraint between nodes  $i$  and  $j$ , with labels  $\lambda$  and  $\lambda'$ , respectively [9].

For the stereo correspondence problem, a simplified version of the above algorithm is used. Let  $I$  be the set of nodes  $i = (x_l, y_l, x_r, y_r, \theta_l, \theta_r, \kappa_l, \kappa_r)$ . Each node is assigned two labels  $\lambda \in \{TRUE, FALSE\}$ . Since there are only two possible labels at each node and  $p_i(TRUE) + p_i(FALSE) = 1$ , we only store  $p_i = p_i(TRUE)$ . ( $p_i(FALSE) = 1 - p_i$ ). The update rule is:

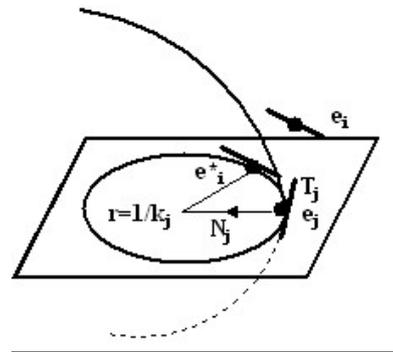


Figure 7: Illustration of transport distance. The nearby tangent at  $e_i$  is projected onto the Frenet approximation computed at  $e_j$ . According to the minimum torsion constraint, this is approximated by projection onto the osculating circle at  $e_j$ , which lies in the osculating plane given by the  $(T_j, N_j)$  frame. Thus three components contribute to the compatibility: transport distance along the osculating circle; projection distance onto the osculating circle; and angular rotation to tangency.

$$p_i^{t+1} = [p_i^t + \sum_j r_{ij} p_j^t]_0^1 \quad (6)$$

where  $[\cdot]_0^1$  denotes projection onto  $[0,1]$ .

The relaxation labeling graph is built from all possible stereo corresponding pairs, as described below.

### 3.3. The Discrete Tangent Map

Edges are detected in each image to provide  $(x, y, \theta, \kappa)$ , with  $(x, y)$  the position coordinates,  $\theta$  the orientation of the (edge) tangent, and  $\kappa$  the curvature. Again following biology, the edge measurements are taken redundantly, one at each orientation, so that multiple tangents can be represented at each image point. This is important for the problem of partial occlusion at junction points. To illustrate: at the image junction in Fig. 1(left), the tangent map has two different tangents, one for the branch in front and one for the branch behind. These two tangents can match with the tangents in the right image along the epipolar lines. As a result, both the branch in front and the branch behind have a valid correspondence pair at the junction point.

To get the discrete tangent map (Fig. 1), we use the logical/linear operators [10] on the original images, followed by a trace inference stage [14] and an interpolation function [4]. The initial edge detection is performed at 16 orientations, and interpolated to 90 quantized orientations and 21 discrete curvatures. The additional accuracy in orientation is required for the orientation disparity component of the compatibility distance function.

### 3.4. Building Compatibility Fields with Near/Far Inhibition

The (left,right) tangent maps provide the candidate matches to build the relaxation network. For each item  $(x_l, y_l, \theta_l, \kappa_l)$  in the left tangent map, a search along the epipolar line in the right tangent map reveals pairs of possible correspondences. Each stereo tangent pair  $(x_l, y_l, x_r, y_r, \theta_l, \theta_r, \kappa_l, \kappa_r)$  is then a node in the stereo relaxation labeling graph. The edge relationship in the graph is given by spatial neighborhoods.

For each stereo tangent pair, we compute its local space curve structure in the osculating plane, as described previously. Fig. 8 illustrates this idea for two corresponding pairs.

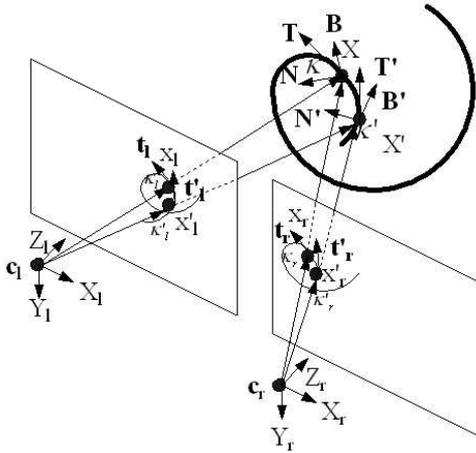


Figure 8: Are two neighboring candidate correspondence pairs compatible? The relationship has to be derived from the geometry of space curve and its image projection curves. Shown are two Frenet frames in  $R^3$  at nearby points along a space curve, which project to two frames in the (left,right) image pair.

Even with the local geometric relationship derived above, however, there are still ambiguities along the epipolar line. To illustrate, consider the highlighted upper right region in Fig. 1. (Fig. 9) shows the reconstruction based solely on the geometric relationships. In the front view of the reconstruction (Fig. 9(c)), the left and right parts have similar - correct - depth, while the middle parts are false matches. This is even more clear in the rotated reconstruction view (Fig. 9(d)), where the correct matches correspond to the middle two branches with similar depth, and the outer two branches are the false matches. Without global constraint, such false matches will generically arise because one point in the left image could possibly match to multiple points in the right image.

This correspondence ambiguity is classical [12], but cannot be dealt with uniformly by the standard heuristics that have been introduced for physical surfaces. Since many

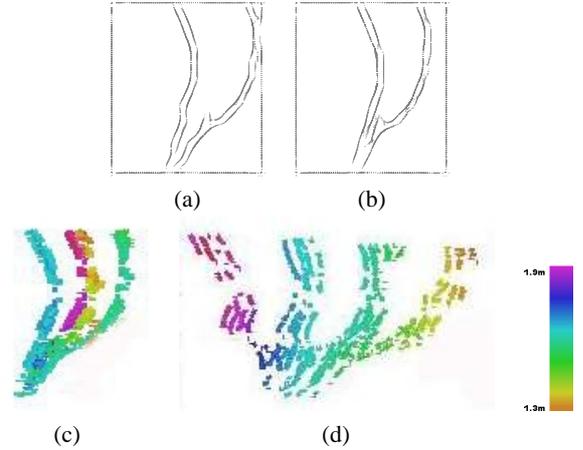


Figure 9: False matches from attempting correspondence on a pair of twigs, because the geometric compatibilities can apply to all possible matches. (E.g., the left branch in (a) can match both the left and right branches in (b), and so on.) (a)(b) Left and right tangent maps of the highlighted upper right region in Fig. 1. (c) Front view of reconstruction showing false matches (the middle two). (d) Rotated view showing false matches plus ghosts from noise in the tangent map. Depth scale shown at right.

of these false matches tend to have large disparity gradients, the *Disparity Gradient constraint* from the PMF algorithm [16] might be appropriate. Neurobiology suggests another implementation of this, which is natural because it fits directly into our compatibility functions.

We remarked earlier that binocular neurons were orientation selective, and exploited that property. We now exploit the fact that they can be classified into three groups that relate to disparity offsets in receptive field structure [15]. (1) *Tuned excitatory neurons* form a group that are disparity selective over a limited (and often narrow) range. (2) *Far neurons* exhibit a selectivity for uncrossed disparities; and (3) *Near neurons* are selective for crossed disparities. Far and near neurons are complementary: One set gives excitatory responses to objects farther than the point of fixation and inhibitory responses to nearer objects; while other set has the opposite behavior, excitation for nearer objects and inhibition for farther ones [11]. It is precisely this property that we include in our compatibility functions, with inhibition for each possible match in proportion to the probability of match ( $p_i$ ); see Fig. 10. The result successfully eliminates the great majority of false matches, as we now show.

## 4. Results and Conclusion

Fig. 11 shows the 3D reconstruction of image pair (Fig. 1) by our algorithm, together with an enlargement of the problem area discussed previously. The color depth map indicates both the front and the back branches are correctly reconstructed. Furthermore, note how the false matches were

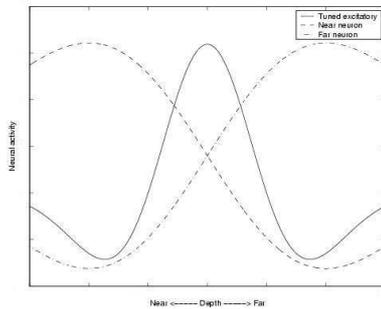


Figure 10: Near and Far neuron tuning function. In biological terms depth can be encoded by neurons tuned to specific disparities; this is cartooned here as an excitatory neuron tuned to 0 disparity. Also shown are tuning curves for near and far neurons, which are excited (resp. inhibited) by stimuli closer (resp. further) than the point of fixation. Our compatibility functions are the product of these tuning curves with the geometric ones (Fig. 5, bottom). Thus potential matches are supported by geometric consistency but inhibited by better matches in the neighborhood.

removed by the near/far inhibition from the upper-right portion (compare Fig. 9). The relaxation process converged after 5 iterations in 60 sec. on a Pentium III PC.

Lacking ground truth on the twig example, we assessed the accuracy of our algorithm with synthetic images. Ten algebraic space curves were generated with parameters chosen randomly. The curves were all started at distance of 1800mm, and ranged in depth from about 1640mm to 1960mm. These were then projected through the camera model to create an image pair, and our algorithm was run on this pair. The results were then compared with ground truth. In Fig. 12(LEFT) we show one example (an ellipse) in which the actual and reconstructed depth are plotted as a function of arc length. The standard deviation is about 1.665mm from the true value, and the maximal difference is 8.04 mm. The quantization in the system is such that one pixel corresponds to about 9.33 mm (at 1.8m), which shows that our edge detection and interpolation are functioning to sub-pixel accuracy.

Finally, to provide a comparison with other algorithms, in Fig. 13 we show our reconstruction for another, more complex twig scene with disparity maps from traditional algorithms. Neither the refined correlation method [21] (Fig. 13(f)) nor the cooperative algorithm [23](Fig. 13(g)) could be made to perform adequately by varying their internal parameters, although the cooperative algorithm performs much better than the correlation method.

By relating the differential structures at image points (position, tangent, and curvature) in the left and right images with the geometry of the space curve point (position, Frenet frame, and curvature), we show how nearby image correspondence pairs should be consistent with the Frenet approximation of the space curve, and how this compati-

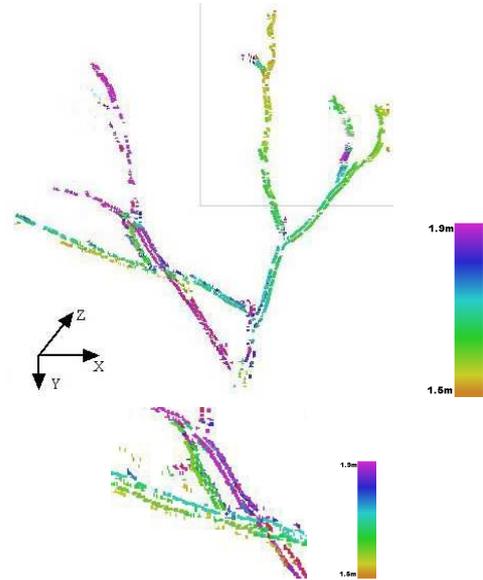


Figure 11: 3D reconstruction of the opening Fig. 1: (TOP) Front view. Note how the false matches were removed by the near/far inhibition (compare Fig. 9). (BOTTOM) A detailed view of the reconstruction of the problem region in which both the ordering constraint and the uniqueness constraint break down. Colored depth scale is shown at right (units: meters).

bility relation can be put into a stereo relaxation labeling network where positional disparity and orientation disparity are combined naturally. Examples demonstrate how our algorithm outperform traditional stereo algorithms, especially at places where heuristic constraints break down.

## References

- [1] S. Alibhai and S. W. Zucker. Contour-based correspondence for stereo. In *Proc. 6th ECCV*, 2000.
- [2] R. Cipolla and P. Giblin. *Visual Motion of Curves and Surfaces*. Cambridge University Press, 2000.
- [3] R. Cipolla and A. Zisserman. Qualitative surface shape from deformation of image curves. *IJCV*, 8:53–69, 1992.
- [4] C. David and S. W. Zucker. Potentials, valleys, and dynamic global coverings. *IJCV*, 5:219–238, 1990.
- [5] O. Faugeras. *Three-Dimensional Computer Vision*. The MIT Press, 1993.
- [6] O. Faugeras and L. Robert. What can two images tell us about a third one? *IJCV*, 18:5–19, 1996.
- [7] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [8] I. P. Howard and B. J. Rogers. *Binocular Vision and Stereopsis*. Oxford University Press, 1995.
- [9] R. A. Hummel and S. W. Zucker. On the foundations of relaxation labeling processes. *IEEE Trans. on PAMI*, 5(3):267–287, 1983.

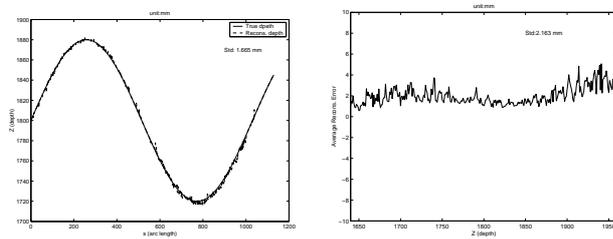


Figure 12: Error analysis from a series of 10 algebraic curves with random parameters (see text). (LEFT) Plot of reconstructed depth and true depth vs. arc length for an ellipse; standard deviation: 1.665mm. (RIGHT) Average absolute value of the difference between the reconstructed depth and the true depth, vs. depth. Standard deviation: 2.163mm.

[10] L. A. Iverson and S. W. Zucker. Logical/linear operators for image curves. *IEEE Trans. on PAMI*, 17(10):982–996, 1995.

[11] S. Lehky and T. Sejnowski. Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity. *J. Neurosci.*, 10:2281–99, 1990.

[12] D. Marr. *Vision*. W.H. Freeman and Company, 1982.

[13] G. Medioni and R. Nevatia. Segment-based stereo matching. *CVGIP*, 31(1):2–18, July 1985.

[14] P. Parent and S. W. Zucker. Trace inference, curvature consistency, and curve detection. *IEEE Trans. on PAMI*, 11(8):823–839, 1989.

[15] G. F. Poggio and B. Fischer. Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *J. Neurophys.*, 40:1392–1405, 1977.

[16] S. B. Pollard, J. E. W. Mayhew, and J. P. Frisby. Pmf: A stereo correspondence algorithm using a disparity gradient limit. *Perception*, 14:449–470, 1985.

[17] L. Robert and O. Faugeras. Curve-based stereo: Figural continuity and curvature. In *Proc. IEEE conf. on CVPR*, 1991.

[18] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47(1/2/3):7–42, 2002.

[19] C. Schmid and A. Zisserman. The geometry and matching of lines and curves over multiple views. *IJCV*, 40(3):199–233, 2000.

[20] Y. Shan and Z. Zhang. New measurements and corner-guidance for curve matching with probabilistic relaxation. *IJCV*, 46(2):157–171, 2002.

[21] C. Sun. Fast stereo matching using rectangular subregioning and 3-d maximum-surface techniques. *IJCV*, 47(1/2/3):99–117, 2002.

[22] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. on PAMI*, 22(11):1330–1334, 2000.

[23] C. Zitnick and T. Kanade. A cooperative algorithm for stereo matching and occlusion detection. *IEEE Trans. on PAMI*, 22(7):675–684, 2000.

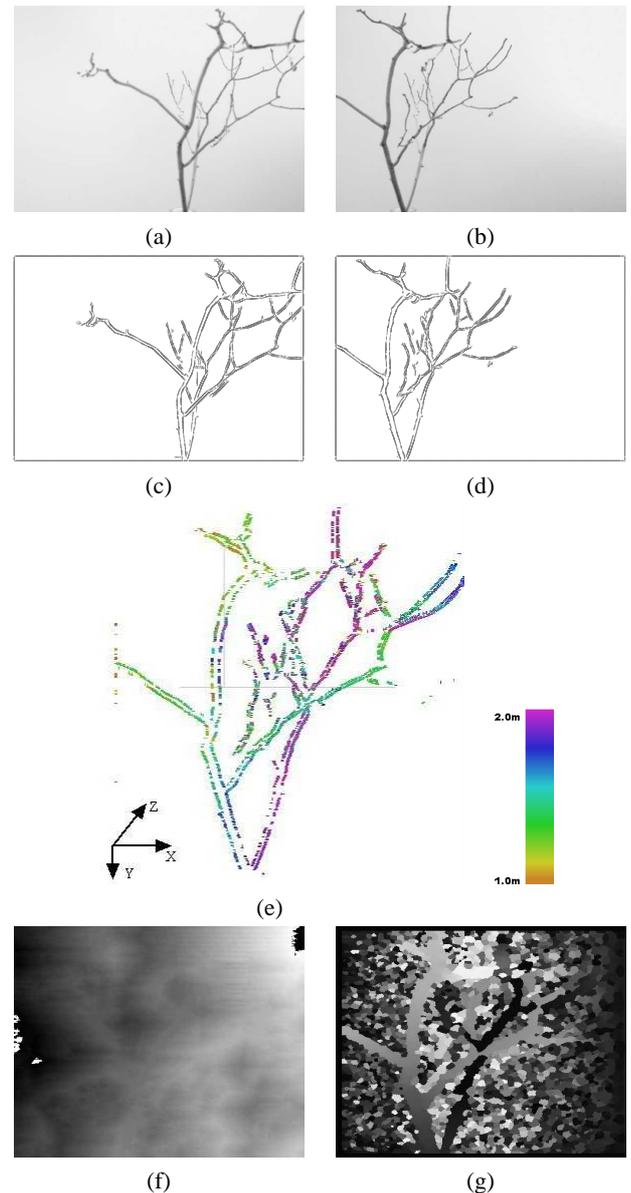


Figure 13: 3D Reconstruction: (a)(b) Image pair. (c)(d) Discrete tangent maps. (e) Front view of reconstruction, colored depth scale shown at right. (f)(g) Disparity maps of traditional algorithms [21, 23].