# Stereo Without Search

Carlo Tomasi and Roberto Manduchi

Computer Science Department
Stanford University, Stanford, CA 94305

**Abstract.** In its traditional formulation, stereo correspondence involves both searching and selecting. Given a feature in one scanline, the corresponding scanline in the other image is searched for the positions of similar features. Often more than one candidate is found, and the correct one must be selected. The problem of selection is unavoidable because different features look similar to each other. Search, on the other hand, is not inherent in the correspondence problem. We propose a representation of scanlines, called *intrinsic curves*, that avoids search over different disparities. The idea is to represent scanlines by means of local descriptor vectors, without regard for where in the image a descriptor is computed, but without losing information about the contiguity of image points. In fact, intrinsic curves are the paths that the descriptor vector traverses as an image scanline is traversed from left to right. Because the path in the space of descriptors ignores image position, intrinsic curves are invariant with respect to disparity under ideal circumstances. Establishing stereo correspondences is then reduced to the selection of one among few match candidates, a task simplified by the contiguity information carried by intrinsic curves. We analyze intrinsic curves both theoretically and for real images in the presence of noise, brightness bias, contrast fluctuations, and moderate geometric distortion. We report preliminary experiments.

## 1 Introduction

The computation of stereo correspondences has traditionally been associated with a search over all possible disparities: for every point in the left scanline the corresponding right scanline is searched for a similar point. In this paper we show that search over disparities is not inherent in the correspondence problem. The way out of search is associative memory, and essentially inverts the way images are represented. Rather than storing image intensities by their position in the scanline, the usual array $I(x)$, we can store scanline positions by their appearance: in a sense, $x(I)$. Then, image points that look similar are stored in the same place. If both scanlines are stored in the same memory, correspondences are trivially established, because corresponding points share the same memory locations. Occlusions are also easily found as points that live alone in some location. There are two problems with this scheme: ambiguity and disguise.

[0] This research was supported by the National Science Foundation under contract IRI-9496205.

Ambiguity means that different image points can look the same, so memory locations can be crowded, and one match must be selected among many. Disguise occurs when corresponding points in the two scanlines look different because of the viewpoint change or because of image noise. In this case, points that should go in the same memory location do not. We deal with disguise by analyzing possible changes between scanlines. This analysis tells us where to look next if a memory location is missing a point. Ambiguity is addressed by a twofold strategy. On the one hand, it is reduced by encoding image appearance with descriptors that are richer than the mere image intensity $I$: each image location is described by a whole vector of parameters. On the other hand, the resolution of the remaining ambiguity is made easier by preserving contiguity information with the descriptors. Consider traversing a scanline in one of the two images. The vector of descriptors traces a curve in some space, and points that are nearby in the scanline are also nearby in the representation. Contiguity then helps selecting among similar match candidates: when two points look similar, we look around them, and attempt matching entire curve segments at once, rather than isolated points.

To illustrate the approach, here is one simple version of intrinsic curve for, say, the left scanline. A lowpass filtered version of the image intensity $l(x)$ and its derivative $l'(x)$ are computed everywhere (solid lines in figure 1 (b) and (c)) and are plotted against each other (solid lines in figures 2 (a) and (b)). When plotting $l'$ versus $l$ we lose track of space, that is, of the coordinate $x$ which merely parameterizes the curve $l'(l)$. This parameter is stored for later use, but it plays no role in the shape of the curve. If $l(x)$ is replaced by a shifted replica $r(x) = l(x + d)$, the curve of figure 2 (b) remains the same. Because of this invariance to displacements, we call the curve of figure 2 (b) an *intrinsic curve*. More general geometric transformations $r(x) = l(\alpha(x))$ between $l$ and $r$ can deform an intrinsic curve, but the deformations can be predicted. The dashed curves in figures 1 and 2 show the construction of the intrinsic curve for the scanline $r(x)$ taken from a different viewing position.

Ambiguities cause intrinsic curves to self-intersect or overlap, and cannot be avoided. The selection process just mentioned is therefore unavoidable. On the other hand, the richer the description is, that is, the higher the dimensionality in which an intrinsic curve lives, the less likely self-intersections are.

In the next section, we present a theory of intrinsic curves. Section 3 show how real images differ from the ideal case, and section 4 shows preliminary experiments.

## 2 Intrinsic Curves: Theory

An efficient procedure for matching two signals is to consider a vectorial description of the local intensity variation at every point. Then two points from the two images are match candidates if the local descriptions are "close" to each other. A similar idea is at the basis of the stereo algorithms of Kass [14], Jones and Malik [11], and Weng, Ahuja and Huang [30]. In this section we define intrinsic

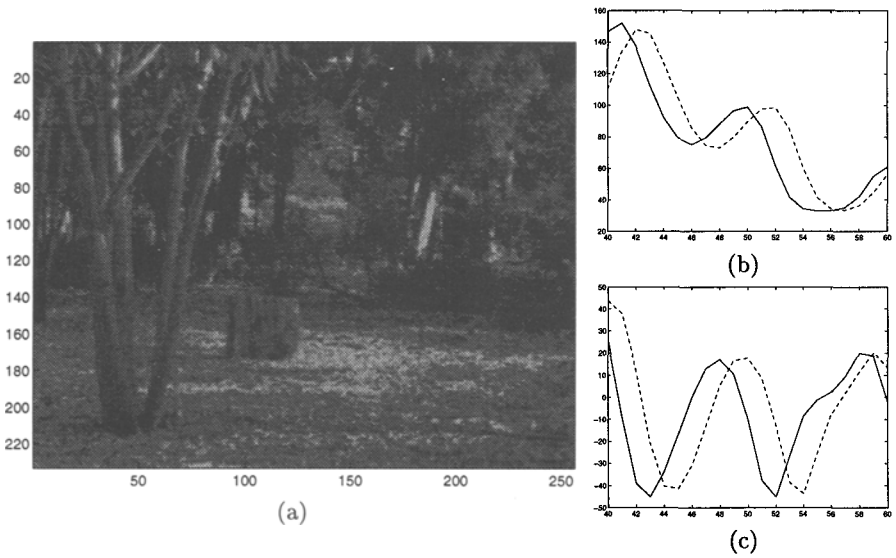



**Fig. 1.** (a) Test image "Trees" from SRI - frame 1. (b) Lowpass-filtered scanline 68, pixels 40–60 (solid line: frame 1, dashed line: frame 2), and (c) its derivative.

curves more generally. We also identify the geometric mappings

$$r(x) = l(\alpha(x)) \tag{1}$$

between the two images that are *compatible* with any particular way of building intrinsic curves, in the sense that they leave the curves unaltered. In other words, intrinsic curves are invariant with respect to compatible mappings. Finally, we investigate geometrical and topological properties of intrinsic curves.

*Definition* of an *intrinsic curve.* Suppose that the $N$ operators $P_1, \ldots, P_N$ are applied to the intensity signal $l(x)$ to produce the new signals $p_n(x) = [P_n l](x)$ for $n = 1, \ldots, N$. The vector

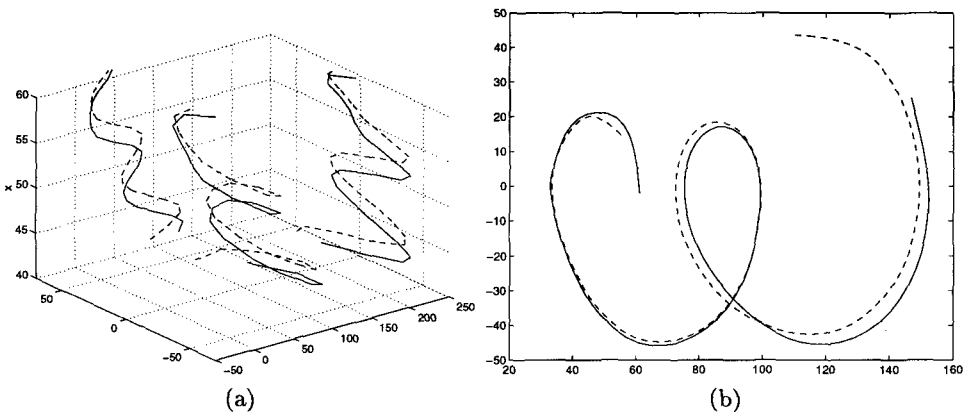
$$\mathbf{p}(x) = (p_1(x), \ldots, p_N(x)) \tag{2}$$

describes a curve $\mathcal{C}$ in $R^N$ parameterized by the real variable $x$:

$$\mathcal{C} = \{\mathbf{p}(x),\, x \in R\}\ . \tag{3}$$

$\mathcal{C}$ is called the *intrinsic curve* generated by $l(x)$ through the operators $P_1, \ldots, P_N$ (see figures 1, 2).

It is crucial to notice that the curve $\mathcal{C}$ lives in $R^N$, not $R^{N+1}$: the image coordinate $x$ is not a component of the curve. Spatial information is lost when going from image $l$ to curve $\mathcal{C}$, and it is exactly this loss of information that makes $\mathcal{C}$ invariant to a suitable class of geometric transformations, as discussed in the following subsection.




(a) (b)

**Fig. 2.** (a) Intrinsic curve formation: the signals of figure 1 (b) and (c) are plotted against each other, forming a 3-D curve whose projection on the plane $x = 0$ is the intrinsic curve (b).

## 2.1 Compatible Mappings

While any reparametrization of $C$ leaves $C$ unchanged, reparametrizing the generator signal $l(x)$ to $l(\alpha(x))$ can in general modify $C$. For instance, if $p_1(x) = l(x)$ and $p_2(x) = l'(x)$, where the prime denotes differentiation, the new components of $C$ after the change $x \to \alpha(x)$ become $\tilde{p}_1(x) = l(\alpha(x))$ and $\tilde{p}_2(x) = \alpha'(x)l'(\alpha(x))$ so that $\tilde{\mathbf{p}}(x)$ traces a new curve that is modulated by $\alpha'(x)$ in its second component. It therefore makes sense to ask what reparametrizations $x \to \alpha(x)$ leave $C$ unaltered.

*Definition* . A mapping $x \to \alpha(x)$ is said to be *compatible* with the operators $P_1, \ldots, P_N$ if for any signal $l(x)$ the intrinsic curve generated by $l(x)$ is equal to the intrinsic curve generated by $l(\alpha(x))$.

### Examples

*Constant Displacement.* Let the operators $P_n$ in (2) be shift-invariant: $l(x) \to l(x + d) \Rightarrow p_n(x) \to p_n(x + d)$. The constant displacements $\alpha(x) = x + d$ are compatible with shift-invariant operators.

*Affine Mapping.* The affine mappings of the form $\alpha(x) = ax + d$ are compatible with the operators $p_0(x) = l(x)$ and

$$p_n(x) = [P_n l](x) = \frac{\left(\frac{\mathrm{d}^n}{\mathrm{d}x^n} l(x)\right)^{(n+1)/n}}{\frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}} l(x)}$$

for $n > 0$, defined wherever $\frac{\mathrm{d}^{n+1}}{\mathrm{d}x^{n+1}} l(x) \neq 0$. This is proved immediately by noting that

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}[l(ax+d)] = a^n[\frac{\mathrm{d}^n l}{\mathrm{d}x^n}](ax+d) \; .$$

*Semi-Commutative Mapping.* If the mapping $x \to \alpha(x)$ is regarded as an operator $A$ applied to $l(x)$, that is, $[Al](x) = l(\alpha(x))$, then $\alpha(x)$ is compatible with operators $P_1, \ldots, P_N$ if there is a change-of-variable operator $[Dq](x) = q(\delta(x))$ such that for every $n$ we have $P_n A = D P_n$ where $\delta$ is a diffeomorphism independent of $n$. In fact, in this case, the mapping $x \to \alpha(x)$ simply reparameterizes the intrinsic curve. Both previous examples are special cases of semi-commutative mappings.

Thus, intrinsic curves can be regarded as invariants with respect to the set of compatible mappings, and provide a more general description than "classical" invariants such as function moments [5],[26],[19],[24]. Affine mappings are a popular model for the transformation between the two images of a stereo pair [15],[13],[6], and shift-invariant filters are often used for image descriptors [14],[11],[12],[18],[17]. The fact that in general affine mappings are not compatible with shift-invariant operators is therefore important. This was pointed out in [14], where a clever analysis of the effect of filtering a signal undergoing affine geometrical distortion is carried out. From the results of [14], we can assume that the intrinsic curves are approximately invariant with diffeomorphisms $x \to \alpha(x)$, so long as the supports of the filters' kernels are narrow and $\alpha(x)$ is close to the identity function. In the remainder of this section we assume that the mapping $\alpha(x)$ is a diffeomorphism (which, in particular, implies that it is monotone and continuous). In addition, we assume throughout this paper that both the input signals $l(x)$, $r(x)$ and the operators $P_n$ are continuous, so that the intrinsic curves are connected.

If the transformation between left and right image were just a mapping $\alpha(x)$ compatible with the operators $P_1, \ldots, P_N$, stereo matching would be nearly trivial. In fact, to determine $\alpha(x)$ from the observation of $l(x)$ and of $r(x) = l(\alpha(x))$, the intrinsic curves are first computed from the two signals. For each signal, the parametrization (3) is stored, so that every point on either curve can be traced back to its image coordinate $x$ via table lookup. Because of compatibility, the two intrinsic curves coincide. For every point $\mathbf{p}$ that belongs to both of them, the corresponding image coordinates are a match, with the sole exception of points where the intrinsic curves self-intersect.

### 2.2 Geometrical and Topological Properties of Intrinsic Curves

Our definition of intrinsic curves is quite general. Their properties depend on the characteristics of the operators $\{P_n\}$ in (2). In this section, we concentrate on the case $N = 2$ with the following choice for these operators:

$$p_1(x) = [P_1 l](x) = l(x) \;\text{ and }\; p_2(x) = [P_2 l](x) = l'(x) \; . \tag{4}$$

Vector $\mathbf{p}(x)$ is thus composed by the first two terms of the Taylor expansion of $l(x)$ around $x$, and each point on the intrinsic curve generated by $l(x)$ represents a description of the local behavior of $l(x)$. With this choice, intrinsic curves are defined on a plane, reminiscent of the *phase space* of systems theory [2]. We can define an orientation at $\mathbf{p}$ by computing the unit-norm tangent $\mathbf{t}(\mathbf{p})$ to the curve:

$$\mathbf{t}(\mathbf{p}) = \frac{(l'(x), l''(x))}{\sqrt{(l'(x))^2 + (l''(x))^2}} . \tag{5}$$

The values of $\mathbf{t}(\mathbf{p})$ depend on the position of $\mathbf{p}$ as follows:

- If $\mathbf{p}$ lies in the upper open half-plane, where $l'(x) > 0$, $\mathbf{t}(\mathbf{p})$ assumes values in the right open half-circle $\{(p_1, p_2); p_1^2 + p_2^2 = 1; p_1 > 0\}$. When $\mathbf{p}$ lies on the lower open half-plane, $\mathbf{t}(\mathbf{p})$ is in the left open half-circle.
- If $\mathbf{p}$ lies on the axis of the abscissas, where $l'(x) = 0$, then $\mathbf{t}(\mathbf{p}) = (0, \pm 1)$. In other words, when crossing the axis of the abscissas, the curve forms an angle of $\pm\pi/2$ with it.

Note that if $l'(x) = l''(x) = 0$ (e.g., in a segment where the signal is constant), $\mathbf{p}$ is singular with respect to $x$ [25]. In such a case, the tangent can be defined by continuity. On the other hand, an intrinsic curve can be singular only on the axis of the abscissas. From these rules it follows that *intrinsic curves are naturally oriented clockwise*: they are traversed left-to-right in the upper half-plane and right-to-left in the lower. Furthermore, *any loop must intersect the axis of the abscissas.*

In general, we may consider intrinsic curves $C$ in $R^M$ of the form $\mathbf{p}(x) = \big(l(x), l'(x), l''(x), \ldots, l^{(M)}(x)\big)$. Each point of the curve represents a local description of the function $l(x)$ by an $M$–term Taylor expansion. The topological properties described above apply to the projection of $C$ onto each plane $\big(l^{(n)}, l^{(n+1)}\big)$, *i.e.* to the curve generated by $\mathbf{P}[l](x) = \big(l^{(n)}(x), l^{(n+1)}(x)\big)$.

### 2.3 Intrinsic Curve Reparametrization

Intrinsic curves are continuous curves on the plane or in a space of higher dimension. For computation, on the other hand, intrinsic curves must have a discrete representation. To this end, we now introduce the *arc length parametrization*, which leads to a variable-rate image sampling that emphasizes "busy" parts of the image. We assume hereafter that $l(x)$ has support in the segment $[x_0, x_1]$.

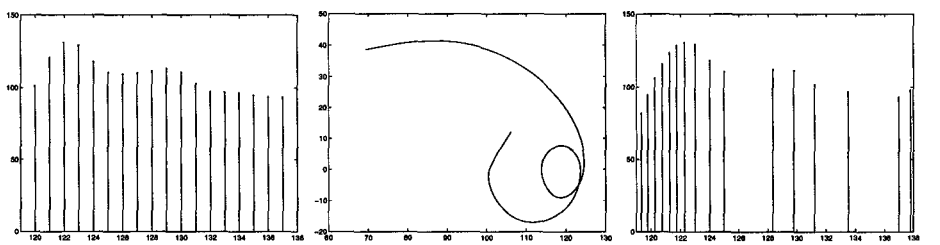
The length of the arc $C(\mathbf{p}_a, \mathbf{p}_b)$ from $\mathbf{p}_a = \mathbf{p}(x_a)$ to $\mathbf{p}_b = \mathbf{p}(x_b)$ is equal to

$$\text{arc length } C(\mathbf{p}_a, \mathbf{p}_b) = \int_{x_a}^{x_b} \sqrt{(l'(x))^2 + (l''(x))^2}\, dx . \tag{6}$$

The arc length parametrization is then $s(x) = \text{arc length } C(\mathbf{p}_0, \mathbf{p}(x))$, where $\mathbf{p}_0 = \mathbf{p}(x_0)$. It is instructive to study the relation between $s(x)$ and $l(x)$. We have that

$$\frac{d}{dx} s(x) = \sqrt{(l'(x))^2 + (l''(x))^2}. \tag{7}$$

Hence, we may expect that a variation $\Delta s$ of the new parameter will correspond to a large variation $\Delta x$ if $l'(x)$ and $l''(x)$ are small (*i.e.*, in parts of the curve that lie close to the horizontal axis), and to a small $\Delta x$ when $l'(x)$ and $l''(x)$ are large (*i.e.*, when the curve is far from the horizontal axis). This observation suggests a sort of "adaptive" sampling paradigm for $l(x)$. Assume to sample the curve $C$ at constant-width intervals, that is, by keeping the arc length of the segments $C(\mathbf{p}_i, \mathbf{p}_{i+1})$ constant. This procedure corresponds to sampling signal $l(x)$ on a nonuniform grid: the grid will be less dense in areas characterized by small values of $l'(x)$ and $l''(x)$ (where the signal is "flat"), and denser if $l'(x)$ and $l''(x)$ are larger (where the signal "busyness" is higher). This looks like a useful sampling strategy for signal matching. In fact, it is well known (see e.g. [9]) that a match is expected to be less robust (with respect, for example, to noise and to quantization errors) in regions where the signal is "flat". The adaptive sampling procedure leads to concentrating estimates in reliable areas(see figure 3).



(a) (b) (c)

**Fig. 3.** A signal sampled on a uniform grid (a) and on an nonuniform grid (c) induced by the uniform arc length sampling of the intrinsic curve (b).

## 3 Deviations from the Ideal Case

Intrinsic curves of corresponding scanlines $l(x)$ and $r(x)$ related by a compatible mapping

$$r(x) = l(\alpha(x)) \tag{8}$$

are identical. In reality, however, $l(x)$ and $r(x)$ can differ for the following reasons.

*No mapping.* In certain cases, $\alpha(x)$ may not even exist, such as when regions of $l(x)$ or $r(x)$ are occluded.

*Incompatible mapping.* The mapping $x \to \alpha(x)$ that relates $l(x)$ and $r(x)$ as in (8) is not compatible with respect to the operators $P_1, \ldots, P_N$ that are used to build the intrinsic curves. For instance, affine transformations $x \to ax + d$ are not compatible with the operator $\mathbf{P}[l](x) = (l(x), l'(x))$. In fact, if $C_l$ is the intrinsic curve generated by $l(x)$, the intrinsic curve $C_r$ generated by $r(x) =$

$l(ax+d)$ is $\mathcal{C}_r = \{(p_1, ap_2) : (p_1, p_2) \in \mathcal{C}_l\}$ which is a vertically expanded ($a > 1$) or compressed ($a < 1$) version of $\mathcal{C}_l$.

*Photometric distortion and noise.* The constant-brightness hypothesis implied by relation (8) is not satisfied. A convenient model that accounts for both geometric and photometric distortion is the following (see [9] for a general discussion of related issues):

$$r(x) = Al(ax + d) + B + n(x) \; . \tag{9}$$

In this model, $A$ and $B$ represent the difference in contrast and brightness between the two images, and are either constant or varying slowly with respect to the dynamic of the signal. The term $n(x)$ represents "noise", that is, any discrepancy independent of the signals. The terms $a$ and $d$ represent geometric distortion and, in particular, $d$ is the inter-frame disparity we are after.

Let us consider the effects of $A$ and $B$ alone (that is, assume $n(x) = 0$ and a compatible $\alpha(x)$). The intrinsic curve $\mathcal{C}_r$ generated by $r(x) = Al(\alpha(x)) + B$ with compatible $\alpha(x)$ is

$$\mathcal{C}_r = \{(Ap_1 + B, Ap_2) : (p_1, p_2) \in \mathcal{C}_l\} \; . \tag{10}$$

Hence, transformation (9) induces an isotropic expansion of the curve by a factor $A$ and a displacement by $B$ along the horizontal direction.

After testing several real-world images, we have observed that the shape of intrinsic curves is altered mostly after photometric distortions, for example as a consequence of the different viewing position of the two cameras, optical attenuation and sensitivity of the image sensors [29], [14], [30]. Large geometric distortions that give raise to vertical dilation or shrinking of the intrinsic curve are less likely to happen than photometric distortions [1]. Consequently, we believe that the terms $A$ and $B$ in our model are dominant over the geometric distortion related to $a$.

The effects of both brightness bias $B$ and noise $n(x)$ can be neutralized by preprocessing both signals with a zero-mean filter with an otherwise lowpass frequency response. The contrast difference term $A$ is then the dominant remaining term, and the point $\mathbf{p}_r$ on $\mathcal{C}_r$ corresponding to a given point $\mathbf{p}_l$ on $\mathcal{C}_l$ is collinear with $\mathbf{p}_l$ and with the origin. Hence, candidates for $\mathbf{p}_r$ are among the points $\{\hat{\mathbf{p}}_r\}$ of $\mathcal{C}_r$ lying on the "radial line" passing through both the origin and $\mathbf{p}_l$. We then select the "right" correspondence within $\{\hat{\mathbf{p}}_r\}$ according to a number of criteria which may be local (proximity in the phase space) or global (ordering and coherence principles). This is described in [28] where we also present a data structure for the efficient access to points that lie along a given radial line.

Our procedure leaves image dilation or shrinking (modeled by the term $a$ in equation (9)) and a possible leftover intensity bias as the only terms of our model (9) that have not been accounted for. An analysis of the inaccuracy due to neglecting such distortion terms may be found in [28].

## 4 A Possible Matching Algorithm

In the ideal case, intrinsic curves from different images coincide, except for occlusions. In reality, because of the phenomena discussed in section 3, intrinsic curves are only close to each other, and matching points can be found along the radial line. A "pathological" case is when the operators $P_n$ are shift-invariant, and $l(x)$ is periodic. In such a case, the intrinsic curve is closed, and infinite instances for $\alpha(x)$ are available. This fact reflects the inherent ambiguity in the match of periodic signals. For non-periodic signals, ambiguity can be somewhat reduced by enriching the description of signals, as noticed also in [14] and [11]. The intrinsic curve representation makes such a notion apparent from a topological standpoint; for example, using only two operators, the intrinsic curves lie in a plane, and self-intersections are to be expected. With three operators, the curves live in a 3-D space, where a path is less likely to cross itself.

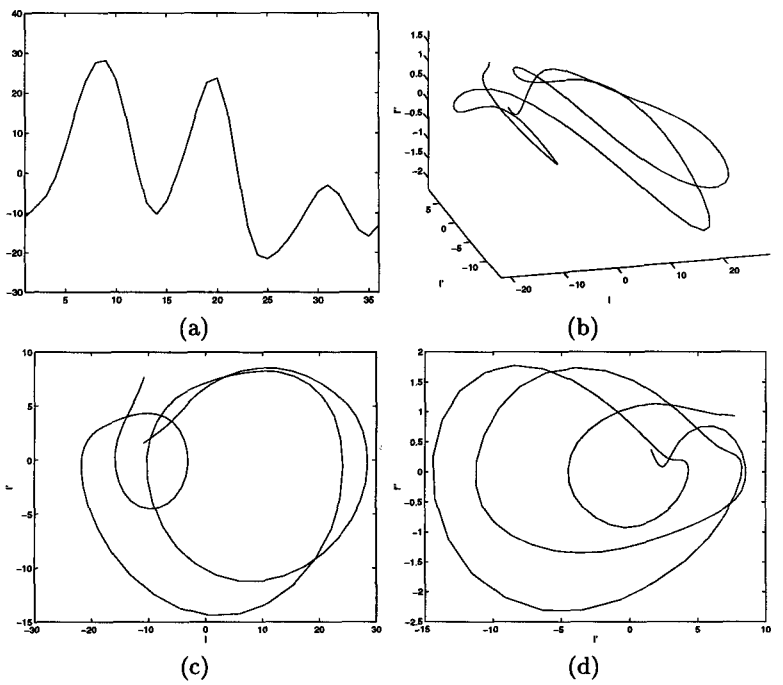



**Fig. 4.** A signal $l(x)$ (a), the intrinsic curve generated by $l(x)$ in the phase space $(l, l', l'')$ (b), and its projections onto the $(l, l')$ plane (c) and onto the $(l', l'')$ plane (d).

Figure 4 shows the intrinsic curve relative to a signal $l(x)$ in the 3-D phase space $(l, l', l'')$, together with its projections on planes $(l, l')$ (our usual intrinsic curve in the plane) and $(l', l'')$. It is interesting to note (from (5)) that the phase of any point of the curve in the plane $(l', l'')$ coincides with the phase of the

tangent to the curve in $(l, l')$ at the corresponding point.

In our algorithm, for each point $\mathbf{p}_l(x)$ in $\mathcal{C}_l$ we select the candidates $\{\hat{\mathbf{p}}_r = \mathbf{p}_r(\hat{x})\}$ of $\mathcal{C}_r$ lying on the radial line passing through $\mathbf{p}_l$ in the plane $(l, l')$. Then, we "rate" each candidate with its distance $d$ to $\mathbf{p}_l(x)$ in the 3-D space $(l, l', l'')$:

$$d(\mathbf{p}_l, \hat{\mathbf{p}}_r) = \sqrt{(l(x) - r(\hat{x}))^2 + (l'(x) - r'(\hat{x}))^2 + (l''(x) - r''(\hat{x}))^2}. \qquad (11)$$

However, the closeness in the phase space alone is not sufficient to guarantee a robust pointwise match of the curves. In fact, approximate ambiguity may occur even in the 3-D phase space when the curve loops close to itself (see figure 4). Hence, to pick the right candidate among $\{\hat{\mathbf{p}}_r\}$, *i.e.*, to resolve the ambiguity, we rely on the available contextual information, as shown in the following.

In the literature, ambiguity is typically resolved by imposing constraints on the disparity field, such as *uniqueness*, *ordering* (or *monotonicity* [7]), and *smoothness* [20],[8], [23],[3],[22],[21]. Note that also other algorithms that make use of vectorial local descriptions ([14], [11]) need to impose constraints on the disparity field: the notion of "closeness" in the representation space is not itself sufficient for a reliable match.

The main novelty of our approach is that *disparity values never enter our procedure to solve the ambiguity.* In fact, we work only on intrinsic curves, which have lost track of space: the inverse mapping $\mathbf{p} \rightarrow x$ is determined only after the matches have been assigned. The new constraint we impose to resolve ambiguities comes naturally from the consideration that under ideal conditions (no photometric distortion, compatible geometric distortion) curves $\mathcal{C}_l$ and $\mathcal{C}_r$ are identical. Let $s$ be the arc length parameter on $\mathcal{C}_l$ (see section 2.3), and let $\Delta s$ be the length of the arc $\mathcal{C}(\mathbf{p}_{l_1}, \mathbf{p}_{l_2})$ between two sampling points $\mathbf{p}_{l_1}, \mathbf{p}_{l_2}$ on $\mathcal{C}_l$. Then, barring occlusions, we expect the length of the (oriented) arc $\mathcal{C}(\mathbf{p}_{r_1}, \mathbf{p}_{r_2})$ on $\mathcal{C}_r$ to be "close" to $\Delta s$. Note that we still rely on the constraints of uniqueness and monotonicity, as parameter $s$ is monotone with $x$, but we do not need any other quality of the disparity field. In other words, we simply expect that, while $\mathbf{p}_l$ moves along $\mathcal{C}_l$, the corresponding point $\mathbf{p}_r$ moves similarly on $\mathcal{C}_r$. We will call such a constraint the *coherence principle.* The important point here is that two corresponding points $\mathbf{p}_l$ and $\mathbf{p}_r$ differ not because of the disparity, but because of noise and distortions. Similarly the arc lengths of $\mathcal{C}(\mathbf{p}_{l_1}, \mathbf{p}_{l_2})$ and $\mathcal{C}(\mathbf{p}_{r_1}, \mathbf{p}_{r_2})$ reflect changes of appearance, not image distances. Details of the algorithm can be found in [28].

For our preliminary experiments, we have chosen three couples of images, namely two frames from the test sequence "Trees" from SRI, two from the sequence "Library", from the movie "Wings of Desire" directed by Wim Wenders, and two from the "Castle" sequence from CMU (figure 5). In all cases, the camera was moving roughly horizontally (*i.e.*, parallel to the scanlines). In the images of figure 5, the part of image above the white line belongs to the first frame, the one below belongs to the second frame.

The left part of sequence "Trees" exhibits a very articulated disparity field, induced by the sharp depth discontinuities along the boundaries of the branches

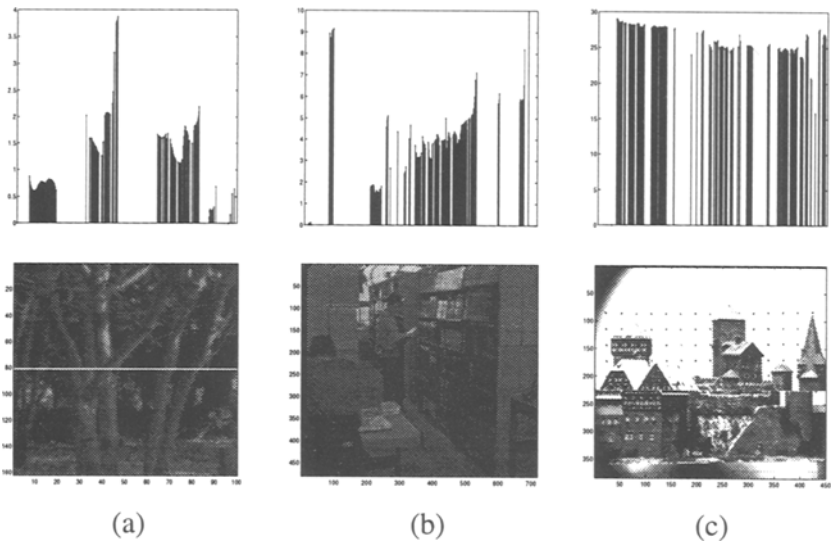



**Fig. 5.** Sequences (a) "Trees", (b) "Library" and (c) "Castle" with the computed disparity fields relative to one scanline (below white line: frame 1, above white line: frame 2).

of the trees. No post–processing (e.g., filtering) has been applied to the results. Note that in some parts we have produced a dense disparity field while in other ones no measurement was available. This is due to the following reasons: (i) the intrinsic curves are sampled with uniform arc length period, which induces the nonuniform sampling period of the measurement field, and (ii) the disparity is computed only where reliable segments of points were found. As pointed out in section 2.3, dense measurements are characteristic of high signal busyness regions, where disparity estimates are more reliable. Figure 5 (a) shows that the computed disparity field follows the depth discontinuities of the scene very tightly. Sequence "Library" (figure 5 (b)) is characterized by a wide disparity range (from less than 2 pixels corresponding to the back of the room, to approximately 10 pixels at the edge of the bookshelves). Both the disparity jump corresponding to the standing person's head (pixels 260–300) and the ramp corresponding to the books on the shelf are detected by our system. The distinguishing feature of sequence "Castle" (figure 5 (c)) is the very large disparity overall: more than 25 pixels. In systems that search over disparities, this large displacement is usually handled by multi–resolution techniques. Our approach, which matches scanlines in the space of descriptors, shows that multi–resolution is not a conceptual necessity.

In these experiments, the variance of the estimates is substantial. However, the measurements are very dense, and a simple post–processing (e.g., median filtering [11]) would "clean" the computed disparity field effectively.

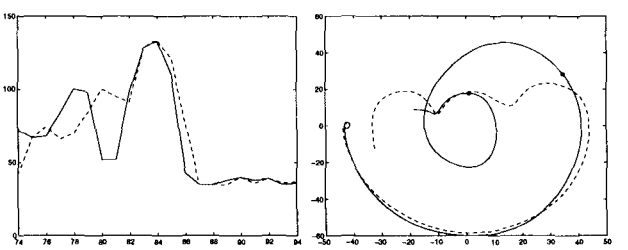
## 5 Future Perspectives

In this paper we have introduced a new image representation which allows approaching the stereo correspondence problem from a new perspective. Our notion of intrinsic curves is a new and useful way to think about stereo, and leads to practical matching algorithms. To the idea of associative storage and retrieval of images, intrinsic curves add the powerful constraint of connectedness. Matching in the space of descriptors makes the amount of disparity irrelevant, and no multi–resolution technique is needed even for large displacements.

Better algorithms can be devised, richer or more stable descriptors can be studied, the robustness to geometric and photometric distortion can be improved. The descriptors can be made even richer through the concepts of local frequency analysis and multi-resolution descriptions, both active areas of research in computer vision today. We hope that the concept of compatible mappings elucidates the basic issues in the design of local image descriptors. Extensions to full images are at the same time conceptually straightforward and technically challenging. The curves become surfaces or manifolds in higher dimensional spaces, but the basis for the matching remains the same.

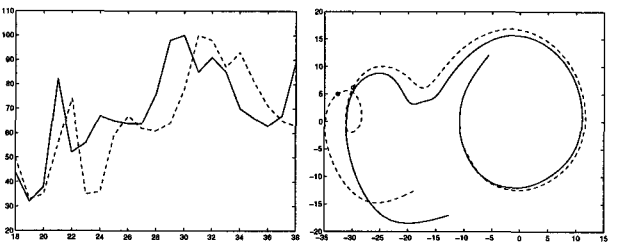
Finally, we would like to outline an intriguing direction of research that we are starting to investigate for the detection of occlusions. Any stereo algorithm must cope with occlusions. A number of researchers have dealt with the problem of occlusions in stereo [27], [16], [11], [4], [7], [10]. The robust and accurate detection of occlusions, however, seems still an open problem.

With intrinsic curves, an occlusion manifests itself as an arc of one curve that is not matched in the other. Just before and just after the unmatched arc the curves are expected to coincide. This situation appears clearly in figures 6(b) and 7(b), where the intrinsic curves of the signals of figures 6(a) and 7(a) are depicted: occlusions stand out as "anomalous" loops in one of the intrinsic curves.



(a) (b)

**Fig. 6.** Scanline 95 from the image of figure 1 (a), pixels # 74–94 (solid line: $l(x)$, dashed line: $r(x)$). (a) Intensity. The part of $l(x)$ form pixel 79 to pixel 81 is not matched by $r(x)$. (b) Intrinsic curves. The arc of $\mathcal{C}_l$ between the two circled points is not matched in $\mathcal{C}_r$.



(a) (b)

**Fig. 7.** Scanline 92 from the image of figure 1 (a), pixels 18–38 (solid line: $l(x)$, dashed line: $r(x)$). (a) Intensity. The part of $r(x)$ form pixel 23 to pixel 25 is not matched by $l(x)$. (b) Intrinsic curves. The arc of $C_r$ between the two circled points is not matched in $C_l$.

In general, the presence of an unmatched loop is not by itself sufficient evidence of occlusion. Loops may be produced sometimes by noise, and we must look for a more robust topological characterization. However, it is clear that an occlusion manifests itself as a "perturbation" of only one of the two intrinsic curves in a limited region. It seems therefore that the detection and analysis of occlusions should be easier in this setting, rather than observing the profiles of the two signals in their "natural" spatial domain. In other words, the phase space is the appropriate place to look whether two signals match – or they don't.

## References


1. R.D. Arnold and T.O. Binford. Geometric constraints in stereo vision. *SPIE* 238:281–292, 1978.
2. V.I. Arnold. *Ordinary Differential Equations.* MIT Press, 1990.
3. H.H. Baker and T.O. Binford. Depth from edge and intensity based stereo. *IJCAI*, 631–636, 1981.
4. P.N. Belhumeur and D. Mumford. A Bayesian treatment of the stereo correspondence problem using half–occluded regions. *CVPR*, 506–512, 1992.
5. A. Blake and C. Marinos. Shape from texture: Estimation, isotropy and moments. *Artificial Intelligence*, 45:323–380, 1990.
6. M. Campani and A. Verri. Motion analysis from first–order properties of optical flow. *CVGIP: Image Understanding*, 56(1):90–107, 1992.
7. D. Geiger, B. Ladendorf, and A. Yuille. Occlusions and binocular stereo. *EECV*, 425–433, 1992.
8. W.E.L. Grimson. Computational experiments with a feature based stereo algorithm. *PAMI*, 7(1):17–34, 1985.
9. W. Förstner. Image Matching. In R.M. Haralick and L.G. Shapiro, *Computer and Robot Vision.* Addison–Wesley, 1992.
10. S.S. Intille and A.F. Bobick. Disparity–space images and large occlusion stereo. *ECCV*, 179–186, 1994.
11. D.G. Jones and J. Malik. A computational framework for determining stereo correspondence from a set of linear spatial filters. *EECV*, 395–410, 1992.

12. D.G. Jones and J. Malik. Determining three–dimensional shape from orientation and spatial frequency disparities. *EECV*, 661–669, 1992.
13. K. Kanatani. Detection of surface orientation and motion from texture by a stereological technique. *Artificial Intelligence*, 23:213–237, 1984.
14. M.H. Kass. Computing stereo correspondence. Master's thesis, M.I.T., 1984.
15. J.J. Koenderink and A.J. Van Doorn. Geometry of binocular vision and a model for stereopsis. *Biological Cybernetics*, 21:29–35, 1976.
16. J.J. Little and W.E. Gillet. Direct evidence for occlusion in stereo and motion. *ECCV*, 336–340, 1990.
17. J. Malik and P. Perona. Preattentive texture discrimination with early vision mechanisms. *JOSA – A*, 7(5):923–932, 1990.
18. S.G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *PAMI*, 11(7):674–693, 1989.
19. R. Manmatha. A framework for recovering affine transforms using points, lines or image brightness. *CVPR*, 141–146, 1994.
20. D. Marr and T. Poggio. Cooperative computation of stereo disparity. *Science*, 194:283–287, 1976.
21. G. Medioni and R. Nevatia. Segment–based stereo matching. *CVGIP*, 31:2–18, 1985.
22. Y. Ohta and T. Kanade. Stereo by intra– and inter–scanline search using dynamic programming. *PAMI*, 7(2):139–154, 1985.
23. S.B. Pollard, J.E. Mayhew, and G.P. Frisby. PMF: A stereo correspondence algorithm using a disparity gradient limit. *Perception*, 14:449–470, 1985.
24. J. Sato and R. Cipolla. Extracting the affine transformation from texture moments. *ECCV*, 165–172, 1994.
25. D.J. Struik. *Lectures on Classical Differential Geometry*. Dover, 1988.
26. B.J. Super and A.C. Bovik. Shape–from–texture by wavelet–based measurement of local spectral moments. *CVPR*, 296–301, 1992.
27. P.S. Toh and A.K. Forrest. Occlusion detection in early vision. *ICCV*, 126–132, 1990.
28. C. Tomasi and R. Manduchi. Stereo without search. Tech. Rep. STAN-CS-TR-95-1543, Stanford, 1995.
29. A. Verri and T. Poggio. Against quantitative optical flow. *ICCV*, 171–180, 1987.
30. J. Weng, N. Ahuja, and T.S. Huang. Matching two perspective views. *PAMI*, 14(8):806–825, 1992.