In the *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, 2010

Minimum length in the tangent bundle as a model for curve completion

Guy Ben-Yosef and Ohad Ben-Shahar
Computer Science Department, Ben-Gurion University
Beer Sheva, Israel
{guybeny,ben-shahar}@cs.bgu.ac.il

Abstract

The phenomenon of visual curve completion, where the visual system completes the missing part (e.g., due to occlusion) between two contour fragments, is a major problem in perceptual organization research. Previous computational approaches for the shape of the completed curve typically follow formal descriptions of desired, image-based perceptual properties (e.g, minimum total curvature, roundedness, etc...). Unfortunately, however, it is difficult to determine such desired properties psychophysically and indeed there is no consensus in the literature for what they should be. Instead, in this paper we suggest to exploit the fact that curve completion occurs in early vision in order to formalize the problem in a space that explicitly abstracts the primary visual cortex. We first argue that a suitable abstraction is the unit tangent bundle $\mathbf{R}^2 \times \mathcal{S}^1$ and then we show that a basic principle of "minimum energy consumption" in this space, namely a minimum length completion, entails desired perceptual properties for the completion in the image plane. We present formal theoretical analysis and numerical solution methods, we show results on natural images and their advantage over existing popular approaches, and we discuss how our theory explains recent findings from the perceptual literature using basic principles only.

1. Introduction and Motivation

Visual completion is a fundamental task in visual systems which facilitates the perception of complete objects from visual fragments. When the object is fragmented due to occlusion, the completion is usually called *amodal*. When the object is illusory and its completed boundary curves are subjective (as happens in the Kanizsa triangle [12]), the completion is known as *modal*. These two completion phenomena have been studied extensively in the perceptual, neuropysiological, and computational vision communities, where one of the main questions investigated is the *shape* of the completed curve. However, unlike the former two domains, the computational pursuit for a mathematical description of the completed shape has been less inspired, and indeed less consistent, with the perceptual evidence. Motivated by the multidisciplinary nature of visual completion research, in this paper we present a new and rigorous mathematical shape completion theory, one which is inspired by neurophysiological and perceptual findings and supports them from basic principles only.

Studies of curve completion usually assume that the completed curve is induced by two oriented line segments (hereinafter the *inducers*). Thus, the curve completion problem can be formulated as follows:

Problem 1 Given the position and orientation of two inducers $p_0 = [x_0, y_0; \theta_0]$ and $p_1 = [x_1, y_1; \theta_1]$ in the image plane, find the curve that passes through the inducers and agrees with perceptual and neurophysiological evidence.

Obviously, as phrased above, the problem is underdetermined in the formal sense, and to solve it mathematically one must derive regularization terms that formalize perceptual and physiological insights. While different regularizations have been proposed, they were inspired more by intuition and mathematical elegance, and less by perceptual findings or neurophysiological principles. In the rest of this section we review some of these computational models, the axiomatic approach which often led to them, and the degree of their perceptual and neurophysiological validity.

1.1. Previous computational work

Among the first to model the shape of completed curves was Ullman[21], who addressed the issue from an *axiomatic* perspective via the formalization of (axiomatic) perceptual characteristics that the completed curve should satisfy:

- *isotropy* - invariance to rigid transformations,
- *smoothness* - continuity of first derivative,
- *total minimum curvature* - integral of curvature along the curve should be as small as possible, and
- *extensibility* - any two arbitrary inducers on a completed curve C should generate the same shape as the shape of the portion of C connecting them.

Ullman suggested the *biarc curve* as the shape that complies with his four axioms. According to this model, the completed curve between two inducers consists of two circular arcs, each osculating to a different inducer and to each other. Since there are infinitely many pairs of such circles, the selected pair is the one that generates the minimal total curvature out of all the admissible pairs. However, despite Ullman's initial intuition, his model was shown to violate extensibility in some cases [3].

After Ullman, many subsequent computational approaches to curve completion have usually followed the same basic axiomatic approach, but possibly with different set of axioms or by putting certain axioms in a more general scope. Chief of all, perhaps, is the axiom of total minimum curvature, which Ullman took in the narrow scope of biarc curves while Horn [10], and later Mumford [16], extended to general regular curves via the model of *elastica* – the family of curves that minimize total curvature

$$\int k(s)^2 ds. \tag{1}$$

The Euler-Lagrange equation applied to this functional 1 leads to a differential equation that must be solved in order to solve for the elastica curve given the inducers. One form of this equation in arclength parametrization is

$$\dot{\theta}^2 = A\sin\theta + B\cos\theta, \tag{2}$$

where θ is the tangential angle of the curve at each position [10]. While this equation requires the determination of two parameters, the most explicit solutions suggested are based on elliptic integrals [10] or theta functions [16]. Unfortunately, no closed-form analytic solution has been found.

Taking the total curvature axiom to its fullest scope comes with a price as it is known that elastica violates *scale invariance* – another axiom that was suggested later (e.g., [23, 18, 14], inspired by Knuth [15]) to argue that the completed shape should be independent on the viewing distance. To make minimum total curvature and scale invariance coexist, a form of a scale invariant elastica was suggested [23, 18], which minimizes the functional $L\int_0^L k^2 ds$, where L is the length of the curve.

Looking for a completion model that captures the nature of "the most pleasing" curve, Kimia et al. [14] proposed the minimization of an energy functional that penalizes *change* in curvature, rather than curvature proper, i.e.,

$$\int \left(\frac{dk}{ds}\right)^2 ds \rightarrow \min \tag{3}$$

subject to the boundary conditions $[x_0, y_0; \theta_0]$ and $[x_1, y_1; \theta_1]$. This model immediately entails a linear expression for the curvature as a function of the arclength, a class of curves known in the mathematical literature as *Euler Spirals*. The Euler spiral model was proved to satisfy all the axioms mentioned above and another popular axiom, the axiom of *roundedness* (again inspired by Knuth [15]), states that the shape of a completed curve induced by two *cocircular* inducers should be a circle. Unfortunately, the scale-invariance and the roundedness axioms were refuted psychophysically in recent perceptual literature, as we will discuss momentarily.

Finally, it should be mentioned that a somewhat different, non axiomatic approach was taken by Williams and Jacobs [26] in their stochastic model to curve completion, which employs assumptions on the *generation process* (in their case, a certain random walk with used-defined parameters) rather than on the desired shape. Although our theory differs from their in both the assumptions and implications, we do share the same philosophy that shape constraints can be *inferred* from more basic principles. Indeed, although it was not verified against perceptual findings, in their work they showed that the stochastic completion field model implies that the most likely completion minimizes certain energy that combines both length and total curvature.

1.2. Relevant perceptual insights

Modern interest in visual completion was triggered primarily by the striking demonstrations of modal completion made by Kanizsa [12]. Besides attempting to model the shape of the completed curve (a.k.a *the shape problem*), perceptual researchers were also interested when two different inducers are indeed *grouped* to induce a curve (a problem often referred to as *the grouping problem*). One popular theory addressing the grouping problem is the *relatability theory* [13], which suggests that two inducers are part of the same contour if their linear extensions intersect in an acute exterior angle. This condition was later shown to parallel the existence of a smooth curve with no inflection points connecting the two inducers [20].

Perceptual studies of the shape problem have been focusing on measuring and characterizing the visually completed curve in different experimental paradigms. For example, the *dot localization paradigm* [9] is a method where observers are asked to localize a point probe either inside or outside an amodally or modally completed boundary, and the distribution of responses is used to determine the likely completed shape. Among the insights that emerged from this and other approaches (e.g., by oriented probe localization [19, 5]) it was concluded that completed curves are perceived (i.e, constructed) quickly, that their shape deviates form constant curvature and thus defies the roundedness axiom (e.g., [9, 19]), and that the completion depends on the distance between the inducers and thus violates scale invariance (e.g., [4, 7]).

1.3. Relevant neurophysiological insights

The famous investigations made by Hubel and Wiesel [11] into the primary visual cortex (V1) have

shown that orientation selective cells exist at complete range of orientations (and at various scales) for all retinal positions, i.e., for each "pixel" in the visual field (Fig 1A). This was captured by the so-called *ice cube* model suggesting that V1 is continuously divided into full-range orientation *hypercolumns*, each associated with a different image (or retinal) position [11]. Hence, an image contour is represented in V1 as an activation pattern of all those cells that correspond to the oriented tangents along the curve's arclength (Fig 1B).

The participation of early visual neurons in the representation of curves is not limited to viewable curves only, and was shown to extend to completed or illusory curves as well. For example, von der Heydt et al. [22] have examined the activity of cells in the visual cortex of Macaque monkeys during presentation of modally completed shapes. They have found that about third of the orientation selective cells in V2 (which are similar in structure to those in V1) fire when presented with such stimuli, and that many of these cells responded nearly the same to real and subjective contours. Grosof et al. [8] found cell responses for illusory contours in about half the neurons studied in V1 of Macaque monkeys, further supporting the conclusion that curve completion is an early visual process that takes place as low as the primary visual cortex.

2. Curve completion in the unit tangent bundle

The basic neurophysiological findings mentioned above suggest that we can abstract orientation hypercolumn as infinitesimally thick "fibers", and place each of them at the position in the image plane that is associated with the hypercolumn. Doing so, one obtains an abstraction of V1 by the space $\mathbf{R}^2 \times \mathcal{S}^1$ [2], as is illustrated in Fig. 1C. This space is an instance of a fundamental construct in modern differential geometry, the *unit tangent bundle* [17] associated with $\mathbf{R}^2$. A tangent bundle of a manifold S is the union of tangent spaces at all points of S. Similarly a unit tangent bundle is the union of unit tangent spaces at all points of S. In our case, it therefore holds [17] that

Definition 1 *Let $I = \mathbf{R}^2$ the image plane. $T(I) \stackrel{\triangle}{=} \mathbf{R}^2 \times \mathcal{S}^1$ is the (unit) tangent bundle of I.*

Recall the aforementioned observation that an image contour is represented in V1 as an activation pattern of all those cells that correspond to the oriented tangents along the curve's arclength. Given the $\mathbf{R}^2 \times \mathcal{S}^1$ abstraction, and remembering that the tangent orientation of a regular curve is a continuous function, we immediately observe that the representation of a regular image curve $\alpha(t)$ is a curve $\beta(t)$ in $T(I)$ (Fig. 1C,D).

We now come to our main idea. If curve completion (like many other visual processes) is an early visual process as early as V1 (cf. Sec. 1.3), and if V1 can be abstracted as the space $T(I)$, then perhaps this completion process should be examined and investigated in *this* space, rather than in the image plane I. In this paper we offer such a mathematical investigation whereby curve completion is carried out in $T(I)$, followed by projection to I. Part of the our motivation for this idea is that unlike the debatable perceptual axioms in the image plane, perhaps the $T(I)$ space, as an abstraction of the cortical machinery, offers more basic completion principles, from which perceptual properties emerge as a consequence.

2.1. Admissibility in $T(I)$

The first question that comes to mind given the above proposal is how curve completion in $T(I)$ would be any different from curve completion in I, namely the image plane. The answer, in fact, is fundamental to our research.

We first observe that when we switch to $T(I)$, Problem 1 turns to deal with the construction of a curve between boundary *points* (e.g., the green points in Fig. 2), rather than between *oriented inducers*. More importantly, as we now discuss, the completed curves in $T(I)$ are not arbitrary and class of curves that we can consider in the first place is quite constrained.

Let $\alpha(t) = [x(t), y(t)]$ be a regular curve in I. Its associated curve in $T(I)$ is created by "lifting" α to $\mathbf{R}^2 \times \mathcal{S}^1$, yielding a curve $\beta(t) = [x(t), y(t), \theta(t)]$, which satisfies

$$\tan\theta(t) = \frac{\dot{y}(t)}{\dot{x}(t)} \quad \dot{x}(t) \equiv \frac{dx}{dt}, \;\; \dot{y}(t) \equiv \frac{dy}{dt}. \qquad (4)$$

We emphasize that $\alpha(t)$ and $\beta(t)$ are intimately linked by Eq. 4, and that α is the projection of β back to I. Examples of such corresponding curves are shown in Fig. 1D.

One can immediately notice that while every image curve can be lifted to $T(I)$, not all curves in $T(I)$ are a lifted version of some image curve. We therefore define

Definition 2 *A curve $\beta(t) = [x(t), y(t), \theta(t)] \in T(I)$ is called admissible if and only if $\exists \alpha(t) = [x(t), y(t)]$ such that Eq. 4 is satisfied.*

There are more inadmissible curves in $T(I)$ than admissible ones, and examples for both curves are shown in Fig. 2. Therefore, any completion mechanism in $T(I)$ is restricted to admissible curves only, and we shall refer to Eq. 4 accordingly as the *admissibility constraint*.

2.2. "Minimum action" completion in $T(I)$

What curve completion principles could we adopt in $T(I)$? In general, since both are vector spaces, any principle that one could use in I is a candidate principle for completion in $T(I)$ also. However, when we recall that $T(I)$ is an abstraction of V1, first completion principle candidate should perhaps attempt to capture likely behavior of neuron populations rather than perceptual criteria (such as

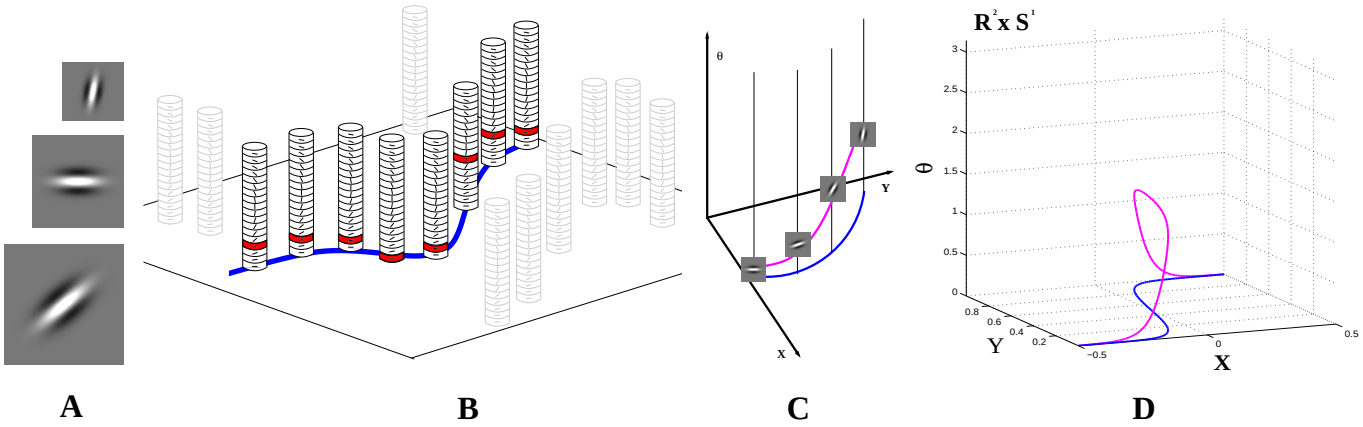

Figure 1. Organization and mechanisms in the primary visual cortex, and its abstraction in the unit tangent bundle. **(A).** Orientation selective cells in V1 come in different scales and orientation tuning. Their receptive fields (shown here is the even-symmetric type only, coded by intensity) often are modeled using Gabor functions or via difference of Gaussians. **(B).** The primary visual cortex is organized in orientation hypercolumns [11], which implies that every retinal position is covered by neurons of all orientations. Thus, for each "pixel" p in the visual field we can think of a vertical vector of orientation-selective cells extending over p and responding selectively (shown in red) according to the stimulus that falls on that "pixel". **(C).** The organization of V1 implies it can be abstracted as the unit tangent bundle $\mathbf{R}^2 \times \mathcal{S}^1$, where vertical fibers are orientation hypercolumns, and the activation patterns to image curves (in blue) becomes "lifted" curves (in magenta). The only aspect of $\mathbf{R}^2 \times \mathcal{S}^1$ not captured in this sketch is the periodicity of it orientation dimension. Thus, the reader should think of the plotted space as quotient space [25] $\mathbf{R}^3/\sim$ where the equivalence structure is such that $(x_1, y_1, z_1) \sim (x_2, y_2, z_2)$ if and only if $x_1 = x_2$, $y_1 = y_2$, and $|z_1 - z_2| = \pi$. **(D).** According to the $\mathbf{R}^2 \times \mathcal{S}^1$ abstraction, any regular curve α (in blue) in the image plane I is represented by a corresponding curve β (in magenta) in $\mathbf{R}^2 \times \mathcal{S}^1$. β can be created by "lifting" α into $T(I)$, such that $\alpha(t)$ and $\beta(t)$ are linked by the admissibility constraint in Eq. 4.

the axioms discussed above). Perhaps the simplest of such principles is a 'minimum energy consumption" or "minimum action" principles, according to which the cortical tissue would attempt to link two boundary points (i.e., active cells) with the minimum number of additional active cells that give rise to the completed curve. In the abstract this becomes the case of the *shortest admissible path* in $T(I)$ connecting two endpoints (x_0, y_0, θ_0) and (x_1, y_1, θ_1). Note that while such a curve in I is necessarily a straight line, most linear curves in $T(I)$ are "inadmissible" in the sense of Definition 2. Since the shortest *admissible* curve in $T(I)$ has a non trivial projection in the image plane, we hypothesize that the "minimum action" principle in $T(I)$ corresponds to the visually completed curve, whose geometrical and perceptual properties are *induced* from first principles rather than *imposed* as axioms.

3. Minimum length curve completion in $T(I)$

Given the motivation, arguments, and insights above, we are now able to define our curve completion problem formally. Let p_0 and p_1 be two given endpoints in $T(I)$ which represent two oriented inducers in the image plane I. The shortest admissible path in $T(I)$ between these two given endpoints is,

$$\ell = \int_{t_0}^{t_1} \sqrt{\dot{\beta}(t)^2} dt \ .$$

For mathematical convenience, however, here we pursue a simplified version of this functional, expressed in the following problem formulation:

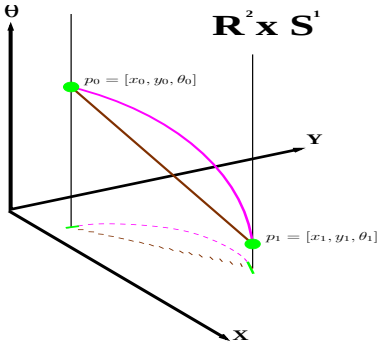

Figure 2. Curve completion and admissible curves in $T(I)$. The completion of two inducers in I is a curve connecting two *endpoints* in $T(I)$. (Both the inducers and endpoints shown here in green). However, regular curves are structures that continue in the direction of their tangent, hence only curves that satisfy the admissibility constraint in Eq. 4 are candidates for completion in $T(I)$ (e.g., the magenta curve and its projected retinal tangents). Unavoidably, *inadmissible* curves in $T(I)$ represent inconsistent pairing of position and tangent orientation in I, and hence cannot represent image curves of any kind (as demonstrated by the "completions" in dark red).

Problem 2 *Given the two endpoints* $p_0 = [x_0, y_0, \theta_0]$ *and* $p_1 = [x_1, y_1, \theta_1]$ *in* $T(I)$*, Find the curve* $\beta(t) = [x(t), y(t), \theta(t)]$ *that minimizes the functional*

$$\ell = \int_{t_0}^{t_1} \dot{\beta}(t)^2 dt = \int_{t_0}^{t_1} [\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{\theta}(t)^2] dt \quad (5)$$

while satisfying the boundary conditions $\beta(t_0) = p_0$ *and* $\beta(t_1) = p_1$*, and the admissibility constraint from Eq. 4.*

Before we solve the constrained variational problem of Problem 2, an important observation can be made already

from its definition. Writing down explicitly our problem as in Eq. 5 reveals an important property of considering curve completion in the unit tangent bundle — it offers a natural and unified substrate for exploring *combinations* of properties of image/retinal curves. Indeed, the sought-after curve in $T(I)$ comprises an approximation of a parameter-free combination of two visual properties in the projected image plane – minimum total curvature ($\int \dot{\theta}(t)^2 dt$) and minimum length ($\int [\dot{x}(t)^2 + \dot{y}(t)^2] dt$). As we discuss in Sec. 4 and the summary, this formal, though elegant and unified way of considering multiple perceptual properties carries over to all types of completion principles that one might care to apply in $T(I)$ – a direct consequence of the explicit representation of both position and orientation at each point in this space. In addition to being directly motivated by the biological visual machinery, this approach becomes particularly appealing in the light of recent findings and conjectures already reported in the perceptual literature, as we shall see below.

3.1. Variational analysis

We now turn to obtain the curves of shortest admissible path between two endpoints in $T(I)$, as defined in Problem 2. At this point we limit the discussion to retinal curves $\alpha(\cdot)$ which are functions over some coordinate system in the image plane, i.e., $\alpha(x) = [x, y(x)]$, as assumed in many previous studies of the curve completion (e.g.,[10, 18]). Consequently, all admissible curves in $T(I)$ are 'lifted' functions of the form

$$\beta(x) \quad = [x, y(x), \theta(x)] \ ,$$

where x becomes the curve parameter (as demonstrated in Figs. 1 and 2). With this description, the functional ℓ from Eq. 5 becomes

$$\ell = \int_{x_0}^{x_1} [1 + \dot{y}(x)^2 + \dot{\theta}(x)^2] dx, \tag{6}$$

and its admissibility constraint (Eq. 4) turns to

$$\tan\theta(x) = \dot{y}(x) \tag{7}$$

Given these observations, the following is our main theoretical result in this paper

Theorem 1 *Admissible curves in T(I) that minimize functional 6 belong to a two parameter family which is defined by the following differential equation*

$$\left(\frac{d\theta}{dx}\right)^2 = \tan^2\theta + A \cdot \tan\theta + B \ . \tag{8}$$

Proof Eq. 7 suggests that functional 6 can be written as a function of θ and its derivative only

$$\ell = \int_{x_0}^{x_1} [1 + \tan^2\theta + \dot{\theta}^2] dx \tag{9}$$

with the boundary condition becoming $\theta(x_0) = \theta_0$ and $\theta(x_1) = \theta_1$. However, by eliminating $y(x)$ from the functional we are at risk of ignoring the boundary conditions for y_0 and y_1, that must be introduced back into the problem by other means. This can be done by adding a constraint on the admissible function $\theta(x)$, so that the induced curve is forced to pass through $[x_0, y_0]$ and $[x_1, y_1]$. Expressed in terms of $\theta(x)$ this constraint becomes:

$$y_1 - y_0 = \int_{x_0}^{x_1} \dot{y} dx = \int_{x_0}^{x_1} [\tan\theta] dx \ ,$$

or

$$\int_{x_0}^{x_1} [\tan\theta - \tfrac{y_1 - y_0}{x_1 - x_0}] dx = 0 \ .$$

This additional constraint can now be incorporated to our functional 6 using some arbitrary Lagrange multiplier λ [1], which results in the following new form for our minimization problem:

$$\int_{x_0}^{x_1} [1 + \tan^2\theta + \dot{\theta}^2 + \lambda(\tan\theta - \frac{y_1 - y_0}{x_1 - x_0})] dx. \tag{10}$$

While we omit the technical details of the derivation, the resulting Euler-Lagrange equation of functional 10 is

$$-2\ddot{\theta} + \frac{\partial(\tan^2\theta + \lambda\tan\theta)}{\partial\theta} = 0 \ .$$

Multiplying the above equation by $\dot{\theta}$ and applying the chain rule[1] we get

$$-\tfrac{d}{dx}(\dot{\theta}^2 - \tan^2\theta - \lambda\tan\theta) = 0.$$

Finally by integrating,

$$\dot{\theta}^2 - \tan^2\theta - \lambda\tan\theta = \mu \tag{11}$$

where μ is an arbitrary constant. ∎

An alternative proof inspired by an approach employed by Horn [10] is presented in the supplemental material.

3.2. Numerical solution and Experimental results

Equation 8 could be solved immediately had we known how to treat the integral

$$x = \pm \int \tfrac{d\theta}{\sqrt{\tan^2\theta + A\tan\theta + B}} \ .$$

Although one may suspect this to take the structure of an elliptic integral, the substitution $\xi = \tan\theta$, leads to an integral of the form

$$x = \pm \int \tfrac{d\xi}{\sqrt{P_6(\xi)}}$$

[1]The only case where multiplication by $\dot{\theta}$ should be avoided occurs when $\dot{\theta} \equiv 0$. We note that this case corresponds to the trivial solution of a straight line for two colinear inducers, a situation that can be detected by other simple means and solved directly.

where $P_6(\xi)$ is a six-order polynomial. It is known that integrals of this form can be expressed as elliptic integrals when the polynomial in the square root is of order three or four ([1], pp.354-359). Unfortunately, this is not the case here and an analytic closed form solution for this integral, if exists, remains an open question for future research.

Since no analytical solution is currently known, we devised a numerical solution, based on nonlinear optimization that seeks the true values of the parameters A and B when two boundary points $p_0 = (x_0, y_0, \theta_0)$ and $p_1 = (x_1, y_1, \theta_1)$ are given. This optimization is based on the observation that one could select parameter values A and B, construct a curve starting from p_0 in a way that obeys the differential Eq. 8, and then evaluate the correctness of the parameters by assessing the error at x_1. The error $E(A, B)$ between the desired and obtained endpoints at x_1 is then used to update the parameters for the next iteration.

More specifically, for each iteration i with a given starting point p_0 and parameter values A and B, we first solve the differential Eq. 8 via Euler's method

$$\begin{aligned} x_{n+1} &\doteq x_n + h \\ \theta_{n+1} &\doteq \theta(x_{n+1}) = \theta(x_n + h) \approx \theta(x_n) + h \cdot \dot{\theta}(x_n) \\ &= \theta_n + h \cdot \sqrt{\tan^2\theta_n + A \cdot \tan\theta_n + B} \\ y_{n+1} &\doteq y(x_{n+1}) = y(x_n + h) \approx y(x_n) + h \cdot \dot{y}(x_n) \\ &= y_n + h \cdot \tan\theta_n \end{aligned}$$

where h is a preselected step size and the error is of order $O(h^2)$. Note that in order to follow the perceptual literature and the general agreement that visually completed curves cannot incorporate inflection points [13, 20, 10, 18], we intentionally prevented $\dot{\theta}$ from changing sign by considering only the positive (or negative) root of Eq. 8 in the implementation of the aforementioned iteration for θ_{n+1}.

The curve $\beta_i(x) = (x, y_i(x), \theta_i(x))$ computed by this step is then evaluated at $x = x_1$ and the error $E(A, B)$ associated with the current value of the parameters is computed by

$$E(A, B) = \|(x_1, y_1, \theta_1) - (x_{end}, y_{end}, \theta_{end})\|.$$

The new values for A and B are then computed by a gradient ascent on $E(A, B)$. We do note that in our implementation the initial guess for A and B was taken via a quick and coarse brute force sampling of the search domain for minimum.

4. Visual properties

We now turn to examine the properties of our curve completion theory in the context of existing perceptual findings and the geometrical axioms reviewed earlier.

Indeed, several properties of our model can be pointed out regarding the six axioms of curve completion mentioned in Sec 1. First, since our solution is not linked to any specific frame, it is trivially isotropic. Second, since it minimizes total arclength in $T(I)$, it must be extensible in that space and hence in the image plane also. Third, since the completed curves can be described by differential Eq. 8, they possess smoothness.

The first axiom where our model departs from prior solutions is the axiom of roundedness since it is easy to confirm that in general the case of constant curvature ($\kappa = \dot{\theta}/\sqrt{1 + \tan^2\theta} = const$) does not satisfy Eq. 8 (apart from a unique case where $A = 0$ and $B = 1$, which corresponds to two cocircular inducers on the *unit* circle). However, given the refutation of roundedness at the perceptual and psychophysical level (as discussed in Sec. 1), we consider this property an advantage rather than a limitation of our model.

One particularly interesting question is how our proposed model behaves with change of scale. Suppose, without loss of generality, that the two inducers are set L units apart on the x axis, i.e., $p_0 = [0, 0, \theta_0]$ to $p_1 = [L, 0, \theta_1]$. As we have discussed, since the completed shape based on functional 6 should minimize the arclength in the tangent bundle (subject to the admissibility constraint), it takes into account both the spatial arclength and the accumulated change of the tangent orientation in the image plane. Hence, we expect that as L increases, and the spatial arclength becomes the dominant component, its minimization would drive the shape closer to a straight line. On the other hand, when L decreases, we expect the relative contribution of $\dot{\theta}$ to the total cost to increase, which would drive the shape away from "straightness". Put differently, we expect our completion model to be scale variant, as recent perceptual and psychophysical studies advocate (see Sec. 1.2).

To show this more formally, we observe that any (i.e., not necessarily minimal) image curve $\alpha_L(x)$ traveling from $p_0 = [0, 0, \theta_0]$ to $p_1 = [L, 0, \theta_1]$ can be written as a scaled version of some other curve $\alpha_1(x)$ traveling from $[0, 0, \theta_0]$ to $[1, 0, \theta_1]$:

$$\alpha_L(x) = L \cdot \alpha_1(x) = [L \cdot x, L \cdot y(x)] \quad , x \in [0, 1]$$

where

$$\alpha_1(x) = [x, y(x)] \quad , x \in [0, 1] \ .$$

Hence, the corresponding curve of $\alpha_L(x)$ in $T(I)$ is

$$\begin{aligned} \beta_L(x) = &\quad [L \cdot x, L \cdot y(x), tan^{-1}(\tfrac{L\dot{y}}{L\dot{x}})] \\ = &\quad [L \cdot x, L \cdot y(x), tan^{-1}(\dot{y})] \\ = &\quad [L \cdot x, L \cdot y(x), \theta(x))] \qquad , x \in [0, 1] \end{aligned}$$

where $[x, y(x), \theta(x)]$ is the corresponding tangent bundle curve of $\alpha_1(x)$. Note now that the value of our functional applied to β_L is

$$\begin{aligned} \ell(\beta_L) &= \int_0^1 [L^2 + L^2\dot{y}^2 + \dot{\theta}^2]dx \\ &= L^2 \int_0^1 [1 + \dot{y}^2]dx + \int_0^1 [\dot{\theta}^2]dx. \end{aligned} \tag{12}$$

The completed curve between the inducers as predicted by our theory should have the shortest admissible path in

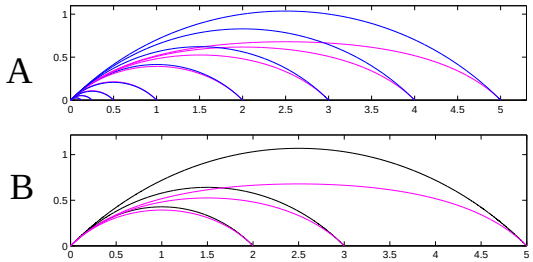

Figure 3. Dependency of the shortest admissible path in $T(I)$ on scale. **(A).** In this example the two inducers are set to $[0, 0, 45^\circ]$ and $[0, L, -45^\circ]$ where $L = 0.125, 0.25, 0.5, 1, 2, 3, 4, 5$. The blue curves are circular completions, which also agree with completions obtained with the Euler spiral model [14]. Shortest admissible paths in $T(I)$ are shown in magenta. Note the "flattening" with increasing scale and "roundedness" with decreasing scale. **(B).** The same shortest admissible curves in $T(I)$ (again shown in magenta) are now compared to the (black) elastica curves [10] for inducers at distances $L \in \{2, 3, 5\}$.

$T(I)$ from p_0 to p_1, and therefore it should minimize functional 12. It is immediately observed that when $L \to \infty$, the curve β_L which minimizes functional 12 approaches the curve that minimizes $\int_0^1 [(1 + \dot{y}^2)]dx$, and thus should be similar to a straight line. On the other hand, when $L \to 0$, the first component of Eq. 12 vanishes, and the resulting minimal curve would minimize $\int_0^1 [\dot{\theta}^2]dx$, subject to the boundary conditions. This entails a curve with approximately linearly changing $\theta(x)$, i.e., a circle-like curve. Fig. 3 demonstrates these two phenomena. and compares the result of our tangent bundle model to both the Euler spiral and the elastica models.

5. Experimental results

In addition to exploring the visual properties as above, we applied our completion model on both synthetic and natural images that incorporate modal and amodal completion. Some results are shown in Fig. 4 and Fig. 5. The natural examples were selected randomly from large data set of natural scenes , and an artificial occluder was added on arbitrarily chosen perceptual contour. Inducers' orientation was measured manually. A canonical scale was used in all images such that distance between adjacent pixels was set to 10^{-2}. Positions and orientations of the two inducers were fed to the numerical algorithm from Sec.3.2 and the parameters A and B were optimized up to an error $E(A, B) \leq 10^{-10}$. The resultant curves of minimum arclength in $T(I)$ were then projected to the image plane I and plotted on the missing parts of the image.

6. Summary and future work

This paper proposes a new theory of curve completion in the unit tangent bundle, the latter being the space which abstracts the early areas in the visual cortex, where curve completion presumably occurs. Employing a universal principle of "minimum action", which strives to minimize neural energy consumption, we have shown that curve completion amount to finding the shortest admissible path in $\mathbf{R}^2 \times \mathcal{S}^1$. We have proved differential properties of this path, showed how it can be found numerically, and derived its perceptual properties.

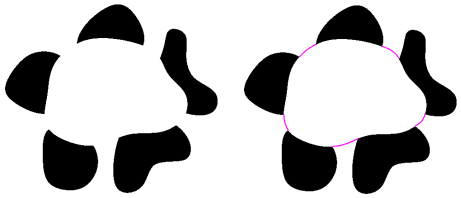

Figure 4. Experimental results for the modal completion. Please zoom in to inspect details.

As we have shown, an implication of the basic "neural energy consumption" constraint suggests parametric-free completions that are determined by minimization of both total curvature and total length in the image plane (see Sec. 3). In this sense our model naturally expresses two basic Gestalt principles [24]. The first, the principle of *good continuation*, is often formalized as minimization of curvature [21]. The second, the principle of *proximity*, is naturally formulated as minimization of total length. While such combinations have been explored in both the perceptual literature (e.g., [4]) and the computational community (e.g., [23, 18, 26]), our theory does so as a result of minimizing a basic (and non visual) principle, from which perceptual insights are derived rather than imposed.

Since, as we also discussed in Sec. 3, the tangent bundle is a substrate for examining *combinations* of image plane properties, it is natural to test additional completion principles in $T(I)$ and explore their implications. For example, one could attempt to study curve of minimum total *curvature* (i.e., elastica) in $T(I)$, which would amount to exploring the combination of both total curvature and total *change* of curvature in the image plane. Interestingly enough, this combination requires boundary conditions about *inducer curvatures* (on top of position and orientation), a type of condition that has been speculated also in the perceptual literature [20, 19].

As illustrated in Fig. 5, the completions our model produces match the desired perceptual outcome and the behavior of natural contours in virtually all cases. Obviously, it is unlikely that all curves in natural images belong to this family of curves. Our result do suggest, however, that an examination of natural contour statistics in the context of contour completion models is an important topic (e.g., [6]), which is another goal of our short term research.

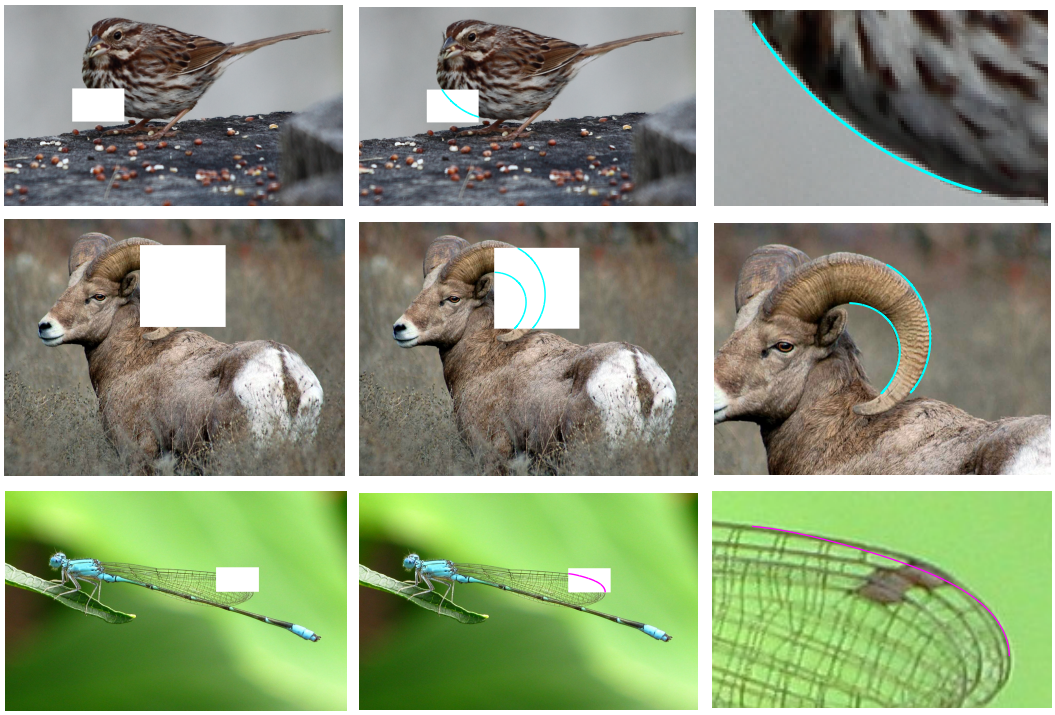

Figure 5. Experimental results of curve completion via shortest admissible curves in $T(I)$ followed by projection to the image plane.

Acknowledgments

This work was funded in part by the Israel Science Foundation (ISF) grant No. 1245/08. We also thank the generous support of the Frankel fund, the Paul Ivanier center for Robotics Research and the Zlotowski Center for Neuroscience at Ben-Gurion University.

References

[1] G. Arfken and H. Weber, editors. *Mathematical Methods for Physicists*. Academic Press, 2001.

[2] O. Ben-Shahar and S. Zucker. The perceptual organization of texture flows: A contextual inference approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(4):401–417, 2003.

[3] M. Brady, W. Grimson, and D. Langridge. Shape encoding and subjective contours. In *Proceedings of the AAAI*, pages 15–17, 1980.

[4] C. Fantoni and W. Gerbino. Contour interpolation by vector-field combination. *Journal of Vision*, 3:281–303, 2003.

[5] J. Fulvio, M. Singh, and L. Maloney. Precision and consistency of contour interpolation. *Vision Research*, 48:831–849, 2008.

[6] W. Geisler and J. Perry. Contour statistics in natural images: Grouping across occlusions. *Visual Neuroscience*, 26:109–121, 2009.

[7] W. Gerbino and C. Fantoni. Visual interpolation in not scale invariant. *Vision Research*, 46:3142–3159, 2006.

[8] D. Grosof, R. Shapley, and M. Hawken. Macaque v1 neurons can signal 'illusory' contours. *Nature*, 365:550–552, 1993.

[9] S. Guttman and P. Kellman. Contour interpolation reveald by a dot localization paradigm. *Vision Research*, 44:1799–1815, 2004.

[10] B. Horn. The curve of least energy. *ACM Transactions on Mathematical Software*, 9(4):441–460, 1983.

[11] D. Hubel and T. Wiesel. Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London, Series B*, 198:1–59, 1977.

[12] G. Kanizsa. *Organization in Vision: Essays on Gestalt Perception*. Praeger Publishers, 1979.

[13] P. Kellman and T. Shipley. A theory of visual interpolation in object perception. *Cognitive Psychology*, 23:141–221, 1991.

[14] B. Kimia, I. Frankel, and A. Popescu. Euler spiral for shape completion. *International Journal of Computer Vision*, 54(1-3):159–182, 2003.

[15] D. Knuth. Mathematical typography. *Bulletin of the American Mathematical Society*, 1:337–372, 1979.

[16] D. Mumford. Elastica in computer vision. In B. Chandrajit, editor, *Algebric Geometry and it's applications*. Springer-Verlag, 1994.

[17] B. O'Neill. *Semi-Riemannian Geometry with applications to relativity*. Academic Press, 1983.

[18] E. Sharon, A. Brandt, and R. Basri. Completion energies and scale. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1117–1131, 2000.

[19] M. Singh and J. Fulvio. Visual extrapolation of contour geometry. *Proceedings of the National Academy of Sciences of the USA*, 102:939–944, 2005.

[20] M. Singh and D. Hoffman. Completing visual contours: The relationship between relatability and minimizing inflections. *Perception & Psychophisics*, 61:943–951, 1999.

[21] S. Ullman. Filling in the gaps: The shape of subjective contours and a model for their creation. *Biological Cybernetics*, 25:1–6, 1976.

[22] R. Von der Heydt, E. Peterhans, and G. Baumgartner. Illusory contours and cortical neuron responses. *Science*, 224:1260–1262, 1984.

[23] I. Weiss. 3d shape representation by contours. *Computer Vision, Graphics and Image Processing*, 41:80–100, 1988.

[24] M. Wertheimer. Untersuchungen zur lehre von der gestalt. *Psychologische Forschung*, 4:301–350, 1923.

[25] S. Willard. *General Topology*. Addison-Wesley Publishing Company, 1970.

[26] L. Williams and D. Jacobs. Stochastic completion fields: A neural model of illusory contour shape and salience. *Neural Computation*, 9(4):837–858, 1997.