SHAPE ENCODING AND SUBJECTIVE CONTOURS*

Mike Brady, W. E. L. Grimson
Artificial Intelligence Laboratory, MIT.
and D. J. Langridge
Division of Computing Research,
CSIRO, Canberra, ACT, Australia.

1. Introduction

Ullman [15] has investigated the shape of subjective contours (see for example [7], [4], [5], [12]). In fact, the work is more generally applicable to other cases of perceptual shape completion, in which the visual system is not constrained by actual physical intensity changes. Examples include patterns formed from dots and incompletely formed line drawings and alphabetical characters.

Ullman proposes that subjective contours consist of two circles which meet smoothly and which are tangential to the contrast boundaries from which they originate. The form of the solution derives from a number of premises, one of which Ullman calls "the locality hypothesis". This is "based in part on experimental observations, and partly on a theoretical consideration" ([15] p2). The "experimental observation" referred to is the following: suppose that A' is a point near A on the filled-in contour AB as shown in Figure 1. If the process by which AB was constructed is applied to $A'B$, it is claimed that it generates the portion of AB between A' and B. Let us call this property "extensibility". Ullman argues that extensibility, together with the properties of isotropy, smoothness, and having minimal integral curvature, logically entails a solution consisting of two circles which meet smoothly. In the full version of this paper, we analyze the two-circle solution and formulate the condition for minimal integral curvature. This can be solved by any descent method such as Newton-Raphson. A program has been written which computes the minimum integral curvature two-circle solution given the boundary angles ϕ, θ, and AB, and which returns as a result the point at which they meet and at which the curvature is discontinuous (the 'knot point'). A tortuous version of this simple proof and program recently appeared in [13]. We then show by example that the two circle solution is not in fact extensible.

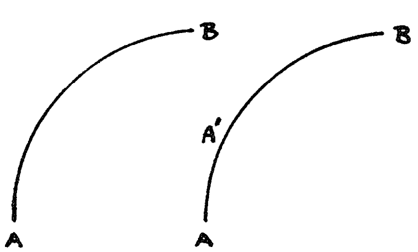

Figure 1. Ullman's extensibility property. If the process by which AB was constructed is applied to $A'B$, it is claimed that it generates the portion of AB between A' and B.

Interestingly, Knuth [8], in his discussion of mathematical typography, initially proposes that "most pleasing curves" should be extensible, isotropic, cyclically symmetric (that is, the solution should not depend on the order of the contributing points), smooth, and be generable by a local (in his case four point) process. In fact, Knuth ([8], Lemma 1) shows that these requirements are mutually inconsistent, and he argues that it is most reasonable to drop the assumption of extensibility. He chooses the mathematically convenient process of cubic spline fitting as the basis of his character typography. This is an interesting choice in the light of what follows here.

We conclude from the above that either Ullman's solution is correct, despite the erroneous justification he gives, in which case we should like to be able to offer an alternative derivation, or else it is incorrect, in which case we should like to propose an alternative.

For a longer discussion of the various suggestions regarding shape completion which have appeared in the literature, the reader is referred to [1]. Often one finds that they are singularly lacking in justification beyond the vague claim that they seem to generate "reasonable" shapes. On a more precise note, Pavlidis([11], chapters 7 and 8) surveys the many versions of polygonal approximation, spline fitting, and Fourier domain descriptors which have been proposed in the considerable pattern recognition literature. Two criteria have tended to dominate the selection of curves for shape completion, namely mathematical tractability and computational efficiency, the latter implicitly assuming some level of computational power on conventional computers. The enormous computational power of the human perceptual system advocates caution in the application of such arguments. Indeed, a somewhat different approach is to base the selection of a class of curves on an analysis of some problem that the human visual system is trying to solve, and to isolate and examine the constraints within which that problem is posed. Several examples of this approach have appeared over the past few years, mainly in the work of Marr and his collaborators (see [9]). The work of Grimson, described briefly in Section 2, is of this sort. Ullman ([16] section 3.3) makes some preliminary comments about the biological feasibility of perceptual computations.

One of the difficulties which surrounds the choice of a class of curves in the specific case of subjective contours is that the differences between alternative choices is often quite small at the angular extent of most examples of subjective contours. This in turn means that the issue is very difficult to resolve by psychophysical experimentation of the sort used by Ullman ([15] page 3).

In [1], we pursue an alternative approach, which was inspired by the second section of Ullman's paper [15]. He develops a local algorithm to compute the two circle solution which minimizes integral absolute curvature. This naturally suggests dispensing with the extensibility assumption entirely, just as Knuth did, and proceeding to find a solution which minimizes some "performance index" related to curvature κ, such as abs(κ) or $\frac{\kappa^2}{2}$. In order to test this idea, we apply the ideas of modern control theory (see [14],[2]).

Since κ and abs(κ) are non-conservative, we consider minimizing $\frac{\kappa^2}{2}$. We develop the Hamiltonian and show that it leads to a particularly intractable differential equation namely,

$$\kappa = \sqrt{2\lambda \frac{dy}{dx} + \nu} \qquad (1)$$

where λ is the Lagrange multiplier, and ν is a constant of integration. This almost certainly does not have a closed form analytical solution. One possible line of approach at this juncture would be to base a local computation on one of the known techniques for the numerical solution of ordinary differential equations. Although shooting methods [3] are an obvious method on which to base such a local computation, we have not yet explored the idea.

In order to proceed, we suggest in Section 2 that the shape completion problem considered here is a two-dimensional analogue of the problem of interpolating a three-dimensional surface, for example from the relatively sparse set of disparity points generated by the Marr-Poggio [10] stereo algorithm. This problem has recently been studied by Grimson [6]. He proposes that the performance index should be a semi-norm and shows that many of the obvious performance indices related to curvature, such as $\frac{\kappa^2}{2}$ and abs(κ), do not have this property. He notes that the quadratic variation, defined by f_{xx}^2 in two dimensions, is not only a semi-norm but is a close approximation to curvature when the gradient f_x is small (subscripts indicate partial derivatives). This is a reasonable condition to impose in the case of subjective contours. Accordingly, we set up the Hamiltonian for the quadratic variation and show that it leads to a cubic, which reduces to a parabola in the case of equal angles. This is particularly interesting in view of the comments made about Knuth's work above.

2. Minimizing quadratic variation

In order to proceed, we suggest that the shape completion problem considered here is a two-dimensional analogue of the problem of interpolating a three-dimensional surface, for example from the relatively sparse set of disparity points generated by the Marr-Poggio [10] stereo algorithm. More generally, we suggest that the process by which subjective contours are generated results from the "mis-application" to two-dimensional figures of a process whose main purpose is the interpolation of three-dimensional surfaces. This idea requires some justification, which space here does not permit (see [1]). It also gives a different perspective on subjective contours, and leads to some fascinating, testable hypotheses.

The three-dimensional interpolation problem has recently been studied by Grimson [6] in the context of human stereopsis. He observes that a prerequisite to finding the optimal function satisfying a given set of boundary conditions (namely that they should all pass through the given set of sparse points), is that the functions be comparable. Translating this into mathematics, he proposes that the performance index should be a semi-norm. Most importantly, he shows that many of the obvious performance indices related to curvature, such as $\frac{\kappa^2}{2}$ and abs(κ), do not have this property. He notes that the quadratic variation, defined by f_{xx}^2 in two dimensions, is not only a semi-norm but is a close approximation to curvature when the gradient f_x is small. This is a reasonable condition to impose in the case of subjective contours, a point which we argue at greater length in [1].

We proceed to set up the Hamiltonian as usual. The boundary conditions are as follows: (1) $x = 0, y = 0, w = \phi$ (2) $x = t, y = 0, w = -\theta = y_1$, where $w = \frac{dy}{dx}$, $y_n = \frac{d^n y}{dx^n}$, and the plant equations are:

$$y_1 = w$$
$$w_1 = u \quad (= y_2)$$

while the performance index is $\frac{u^2}{2}$. The Hamiltonian state function is given by

$$H(u, y, w) = \frac{u^2}{2} + \lambda w + \mu u.$$

Setting $\frac{\partial H}{\partial u}$ equal to zero in the usual way, and solving $\frac{\partial H}{\partial y} = \lambda_1$ and $\frac{\partial H}{\partial w} = -\mu_1$, leads to

$$y_2 = \lambda x - \rho,$$

where ρ is a constant of integration. This integrates easily to yield the cubic solution

$$y = \frac{\lambda x^3}{6} - \frac{\rho x^2}{2} + \sigma x + \tau,$$

where σ and τ are further constants of integration. Inserting the boundary conditions enables us to solve for λ, ρ, σ, and τ. We get finally

$$y = \frac{x^3}{t^2}(\tan\theta - \tan\phi) + \frac{x^2}{t}(\tan\phi - 2\tan\theta) + x\tan\theta.$$

In Figure 2, we showed Ullman's solution for the case $\theta = 30^o, \phi = 20^o$. In Figure 3 we show the curve generated by our method, and in Figure 4 we overlay the two. Clearly, the difference between the solutions is quite small at the angular extent shown in the figures. (Further examples are given in [1]). As we pointed out in the Introduction, this limits the usefulness of the kind of psychophysical experimentation used by Ullman to decide which solution is adopted by humans.

In case the angles ϕ and θ are equal, the cubic term is zero and the solution reduces to a parabola, namely

$$y = \frac{-x^2}{t}\tan\theta + x\tan\theta.$$

The apex is at $\left(\frac{t}{2}, \frac{t}{4}\tan\theta\right)$ and the focal length is $\frac{t}{4\tan\theta}$. Hence the focus is below the line AB so long as $\theta < \frac{\pi}{4}$, which is normally the case for subjective contours.

In the case $\theta = \phi$, it is straightforward to compute the difference between the parabola generated by our method and the circle predicted by Ullman, which is

$$y = \left\{\frac{t^2}{4}\csc^2\theta - \left(x - \frac{t}{2}\right)^2\right\}^{\frac{1}{2}} - \frac{t}{2}\cot\theta.$$

Using the approximation

$$(1-\epsilon^2)^{\frac{1}{2}} \approx 1 - \frac{\epsilon^2}{2} + \frac{\epsilon^4}{8},$$

this reduces to

$$y = \frac{(t-x)}{t} x \tan\theta + \frac{(t-x)^2}{t^3} x^2 \tan^3\theta.$$

The difference between the solutions is essentially given by the second term whose maximum is at $x = \frac{t}{2}$. Dividing by the horizontal extent t of the curve, the relative difference is bounded by $\frac{\tan^3\theta}{16}$. For $\theta < \frac{\pi}{4}$ this is negligible.

3. Acknowledgements

This report describes research done in part at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643. The authors would like to thank the following people for valuable discussions at various stages of the research described here: John Greenman, Pat Hayes, David Marr, Slava Prazdny, Chris Rowbury, Robin Stanton, and Shimon Ullman.

* This paper will appear as MIT AI memo 582. Here intelligibility and content have been sacrificed in the interests of strict page limit brevity.

4. References

[1] **Brady, Grimson, and Langridge** "Shape encoding and subjective contours," (1980), to appear.

[2] **Bryson and Ho** *Applied optimal control*, Ginn, Waltham MA, 1969.

[3] **Conte and de Boor** *Elementary numerical analysis*, McGraw Hill, Tokyo, 1972.

[4] **Coren** "Subjective contours and apparent depth," *Psychol. Review* **79** (1972), 359-367.

[5] **Frisby and Clatworthy** "Illusory contours: curious cases of simultaneous brightness contrast?," *Perception* 4 (1975), 349-357.

[6] **Grimson** Computing shape using a theory of human stereo vision, Ph.D. Thesis, Department of Dept. of Mathematics, MIT, 1980.

[7] **Kanisza** "Subjective contours," *Sci. Amer.* 234 (1976), 48-52.

[8] **Knuth** "Mathematical typography," *Bull. Amer. Math. Soc. (new series)* 1 (1979), 337-372.

[9] **Marr** *Vision*, Freeman, San Francisco, 1980.

[10] **Marr and Poggio** "A theory of human stereo vision," *Proc. R. Soc. Lond. B* 204 (1979), 301-328.

[11] **Pavlidis** *Structural Pattern Recognition*, Springer Verlag, Berlin, 1977.

[12] **Rowbury** Apparent contours, Univ. of Essex, UK,, 1978.

[13] **Rutkowski** "Shape completion," *Computer Graphics and Image Processing* 9 (1979), 89-101.

[14] **Schultz and Melsa** *State functions and linear control systems*, McGraw Hill, New York, 1967.

[15] **Ullman** "Filling the gaps: The shape of subjective contours and a model for their generation," *Biol. Cyb.* 25 (1976), 1-6.

[16] **Ullman** "Relaxation and constrained optimisation by local processes," *Computer Graphics and Image Processing* **10** (1979), 115-125.

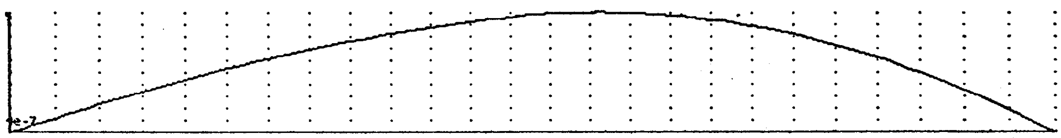

Figure 2. Ullman's solution for the case of boundary angles $\theta = 30^o, \phi = 20^o$.

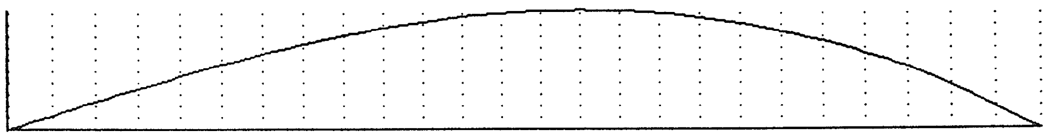

Figure 3. The solution generated by the method proposed in this paper. The boundary conditions are the same as those of Figure 2.

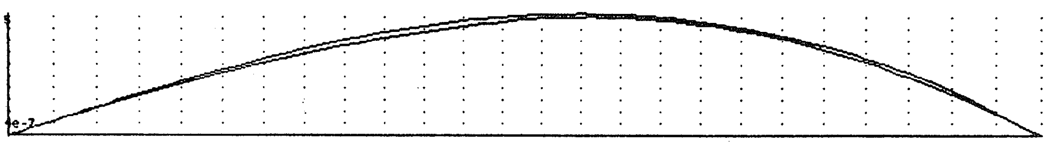

Figure 4. The solutions of Figure 2 and Figure 3 are overlayed to demonstrate the difference between them.