THE SHAPE OF SUBJECTIVE CONTOURS*

Jon A. Webb† and Edward Pervin‡
†Department of Computer Science
Carnegie-Mellon University, Pittsburgh, PA 15217
‡Perq Systems Corporation
Pittsburgh, PA 15217

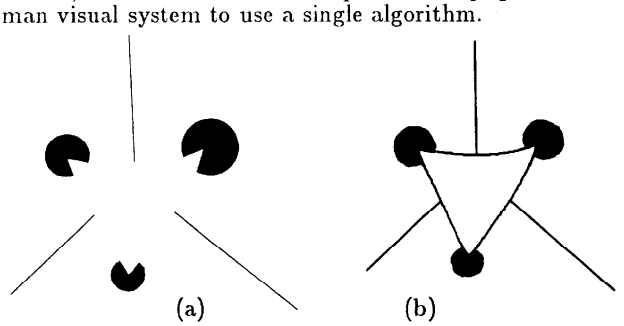
ABSTRACT

We develop a theoretical framework for interpolating visual contours and apply it to subjective contours. The theory is based on the idea of *consistency*: a curve fitting algorithm must give consistent answers when presented with more data consistent with its hypothesis, or the same data under different conditions. Using this assumption, we prove that the subjective contour through two point-tangents is a parabola. We extend the theory to include multiple point-tangents and points. Sample output of programs implementing the theory is provided.

I. INTRODUCTION

Subjective contours are curves filled in by the visual system in the absence of an explicit curve. An example is shown in Figure 1. These curves are relevant to computer vision and graphics, because it is often necessary to fill in missing curves in these fields. If we understand how the human visual system does this, we should be able to program computers to do it.

For this reason, subjective contours have received much attention in computer vision. However, this research has not always been generally useful, because to apply a human visual algorithm to computer vision we must understand more than just the algorithm: we must understand the assumptions on which the algorithm is based, and these assumptions must be precise and referred to the external world, not to other parts of the visual system. Otherwise we would not be able to tell if the algorithm could be expected to work in a specific situation, or we would have to implement large parts of the human visual system to use a single algorithm.



(a) (b)

Figure 1. (a) The Kanizsa triangle [1]. (b) Subjective contour created by our method.

* This research was sponsored by the Defense Advanced Research Projects Agency (DOD) under ARPA Order No. 3597, and monitored by the Air Force Avionics Laboratory under Contract F33615-78-C-1551. The views and conclusions in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U.S. Government.

We will show how a theory of subjective contours can be derived from a few simple assumptions that are referred to the external world. To do this we will start with a few assumptions (principally, that subjective contours are the projections of unseen occluding contours) and show mathematically how this leads to a unique shape for the subjective contour. Knowing the shape we can write programs to draw subjective contours. Because our shape is unique, we have shown that any visual system — human or computer — that makes the same assumptions about contours as we have must produce contours with the same shape.

II. FUNDAMENTAL ASSUMPTIONS.

Consider a visual process that fills in missing contours by connecting scattered data. We assume that this process assumes that the contours it fills in are unseen occluding contours. If we were designing the best possible such process, what might it try to do?

The process cannot always fill in the correct contour because the correct contour is not known, and cannot be predicted. But we will show that the process can be *consistent*: it can give the same answer when presented with more data consistent with what it has seen before, or data equivalent to what it has seen before.

Consistency is important in perception, just as in reasoning. A consistent process can be relied upon to accumulate fragments of data into a complete whole, just as in science we construct a theory by accumulating evidence. An inconsistent process, on the other hand, might arbitrarily change its conclusion as the result of more evidence.

We will now show how consistency leads to specific behavior, which is describable mathematically, and which we can expect from a contour-fitting algorithm.

III. CONSEQUENCES OF CONSISTENCY

The first step in deriving the consequences of consistency is to define it; we can do this as follows. There is a class of transformations on image contours which do not affect the contour seen, but rather represent a change in the position of the viewer: for example, the contour can be rotated, or more of the same contour can be seen. Suppose that the contour-fitting algorithm produces a contour given some scattered data, then we transform the scattered data by one of the transformations above, and we present the transformed data to the contour-fitting algorithm. We will say the algorithm is consistent if it produces the same contour it did originally, but subjected to the transformation.

We will now describe the class of transformations mentioned above. There are two kinds of transformations: one where information is added or replaced, and another where

a set of information is subjected to some visual distortion. The first kind of transformation leads to a criterion called *extensibility*.

A. Extensibility

The principle of *extensibility* is that *adding redundant information to the contour should not affect its shape*. Extensibility is desirable in vision because it makes it possible to deal with an image in which new information is being discovered (perhaps through the action of an on-line visual process). As more information is added to an image, the contours fit through scattered data will change only if the new information is not redundant.

Despite this, most algorithms previously proposed for interpolating subjective contours [3,4,5,6] are not extensible. Extensibility has been rejected primarily because it appears to conflict with other desirable criteria: for example, fitting a smooth contour through the data. But we will show that the contours produced by the extensible algorithm described here are quite smooth and look reasonable. In order to show this, we must precisely define what we mean by extensibility.

To define extensbility precisely, we must introduce some notation. We treat any contour-fitting algorithm as a function $\mathcal{S}(\cdot)$ mapping from a tuple of *contour data* $\langle A_1, A_2, \ldots, A_n\rangle$ to a contour $\mathcal{S}(A_1, \ldots, A_n)$. The contour $\mathcal{S}(A_1, \ldots, A_n)$ passes through the A_i in sequence. Each A_i may indicate either that the contour passes through a point, or that the contour passes through a point with a signed direction. In the second case the A_i is referred to as a *point-tangent*, which is a pair of vectors $\langle P, P'\rangle$. In this pair, P' is a unit tangent vector to the contour at P, and is called the *head* of the point-tangent. The *tail* of the point-tangent is $-P'$. The contour passes through each A_i in the same direction. A subjective contour may be closed or open; if it is closed, we have $A_1 = A_n$. We write $P \in \mathcal{S}(A_1, \ldots, A_n)$ if $\mathcal{S}(A_1, \ldots, A_n)$ passes through the point or point-tangent P. If P precedes Q on the contour (in the sense of the contour direction as defined above) we write $P < Q$.

These definitions make it possible to state extensibility precisely, as an axiom that we assume true of our contour-fitting algorithm:

3.1. Axiom: Extensibility. *If $X \in \mathcal{S}(A_1, \ldots, A_n)$ is a point on a subjective contour such that $A_i < X < A_{i+1}$ then*

$$\mathcal{S}(A_1, A_2, \ldots, A_i, X, A_{i+1}, \ldots, A_n) = \mathcal{S}(A_1, \ldots, A_n)$$

A similar condition applies if $X < A_1$ or $X > A_n$.

Extensibility makes it possible to add data to a contour, but we would also like to be able to replace data on the contour, under certain conditions. We will call this criterion *point replacement*.

B. Point replacement

The ideal contour-fitting algorithm would be able to tolerate arbitrary replacement of data on the contour by other data, but this is too strong a condition to require; if we allowed this, we could move all the data to be nearly adjacent to one point, and it would be unreasonable to expect that we would still get the same subjective contour. We can restrict this criterion while still making it meaningful by allowing only points at the end of the contour to be movable. We can state this restriction as follows:

3.2. Axiom: Point replacement. *If $R \in \mathcal{S}(A_1, \ldots, A_n)$ is a point on a subjective contour such that $R > A_{n-1}$ then $\mathcal{S}(A_1, \ldots, A_{n-1}, R) = \mathcal{S}(A_1, \ldots, A_n)$; similarly if $R < A_2$.*

This completes the development of criteria dealing with the addition and replacement of data. Next we consider criteria arising from viewing transformations.

C. Viewpoint independence

When we shift our point of view we see the same real objects. To be consistent, subjective contours should behave in this way. We can state this principle of *viewpoint independence* as follows. *Changing the point of view, creating a subjective contour, then changing the point of view back should produce the same contour as would be produced from the original point of view.*

We cannot satisfy this condition for all contours; but there is a natural sub-class of contours for which it is satisfiable. We make three different restrictions on the real-world contours that project onto the subjective contours: we restrict the shape of the object along the real-world contour, we restrict the real-world contour's shape, and we restrict its relation to the viewer.

In order to guarantee that the same portion of real contour is seen as we shift our point of view, we require the real contour to be generated by an abrupt bend in the boundary of the object, rather than a smoothly turning boundary. Because the contour-fitting algorithm has no information on the shape of the contour in depth, we require the real-world contour to be planar. Finally, because contours close to the viewer must be modelled using central or point projection, we require the contours to be distant from the viewer, so that we can use parallel or orthographic project, which is mathematically more tractable.

All of these assumptions are quite common in computer vision [7], and they are also common in the real world. For example, leaves have abruptly changing contours that are often planar, and leaves are small enough so that when they are viewed parallel projection is a good model to use.

Planarity, abrupt contours, and parallel projection together imply that subjective contour algorithms must be commutative with *affine transformations* and translations. An affine transformation and translation is a linear transformation $\mathbf{A}: (x, y) \mapsto (ax + by + u, cx + dy + v)$ for some constants a, b, c, d, u, and v. Affine transformations include such things as rotations, skew distortions, and scale changes. Affine commutativity can be stated as follows; note that *any* viewpoint-independent contour-fitting algorithm must satisfy this condition:

3.3. Affine commutativity. *For any subjective contour $\mathcal{S}(P_1, P_2, \ldots, P_n)$, and any affine transformation and translation $\mathbf{A}$,*

$$\mathcal{S}(P_1, \ldots, P_n) = \mathbf{A}^{-1}[\mathcal{S}(\mathbf{A}(P_1), \ldots, \mathbf{A}(P_n))]$$

Now that we have precisely defined all of our restrictions, we can derive the first important result of the paper.

IV. SUBJECTIVE CONTOURS THROUGH TWO POINT-TANGENTS

We will now prove that the subjective contour through two linked point-tangents is either a parabola or two straight lines, where two point-tangents are *linked* if a ray extended from the head of one intersects a ray extended from the tail of the other.

The proof works by first showing that any pair of point-tangents can be transformed into a certain configuration.

Then we show that this configuration can be mapped into itself in a way that produces an infinite number of new points on the curve. All of these points lie on a conic, so that the curve must be a conic. Now from the conditions we have stated it follows that the subjective contour through two point-tangents is a four parameter curve; and since the only four parameter subclasses of the conics are the parabolae and pairs of straight lines, these are the only curves that can be subjective contours.

First we need a lemma dealing with affine transformations:

4.1. Lemma. *For any two pair of non-degenerate linked point-tangents, P, Q and R, S there is an affine transformation and translation $\mathbf{Aff}(P,Q;R,S)$ mapping P onto R and Q onto S. A pair of linked point-tangents $A = \langle U, U'\rangle$ and $B = \langle V, V'\rangle$ is non-degenerate if the triangle formed by U, V, and the intersection point of the head of A with the tail of B is non-degenerate.*

The proof of this is fairly straightforward and is omitted. In the proof of the next theorem, we assume the subjective contour is continuous and connected. Continuity is easy to show for visual contours; since we always see only discretely sampled data, a continuous contour could account for what is seen as well as any discontinous contour. Connectivity can be proved using continuity and our assumptions, but we omit the proof here.

4.2. Theorem. *The subjective contour through any non-degenerate linked two point-tangents is a conic section.*

Proof. Since we can map any pair of non-degenerate linked point tangents onto a given pair, it is sufficient to determine the subjective contour through a specific pair of linked point-tangents, and then to determine the subjective contour through any pair using affine commutativity. Consider two symmetric point-tangents passing through $(1,1)$ and $(-1,1)$. We distort them by an affine transformation that stretches parallel to the x-axis so that the subjective contour between them passes through $(0,0)$. This configuration is invariant to the affine transformation that flips about the y-axis; hence by affine commutativity the tangent at $(0,0)$ is parallel to the x-axis; let it be $(1,0)$. Call the three point-tangents produced in this way $P = \langle(-1,1),(p,q)\rangle$, $Q = \langle(1,1),(p,-q)\rangle$, and $R = \langle(0,0),(0,1)\rangle$. Let the tangent from P intersect the x-axis at $(-k,0)$, so that the tail of Q intersects the x-axis at $(k,0)$.

Consider the affine transformation $\mathbf{A} = \mathbf{Aff}(P,R;R,Q)$. There is no difficulty in writing $\mathbf{A}$ down in terms of k; in homogenous coordinates [8] it is

$$\begin{pmatrix} t & t-1 & 1 \\ t+1 & t & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

where $t = \frac{1}{k} - 1$.

We will now show that $\mathbf{A}$ maps $S(P,Q)$ onto itself.

$$\begin{aligned} S(P,Q) &= S(R,Q) = \mathbf{A}[S(\mathbf{A}^{-1}(R), \mathbf{A}^{-1}(Q))] \\ &= \mathbf{A}[S(P,R)] = \mathbf{A}[S(P,Q)] \end{aligned}$$

Since $S(P,Q) = \mathbf{A}[S(P,Q)]$, it follows that $S(P,Q) = \mathbf{A}^n[S(P,Q)]$, for any n. Now we will consider the effect of $\mathbf{A}$ when repeatedly applied to any point, say R. Since $\mathbf{A}$ maps the subjective contour onto itself, $\mathbf{A}^n(R) \in S(P,Q)$ for any n. Now we shall show that all the $\mathbf{A}^n(R)$ lie on a conic.

We do this by showing that there is a matrix C such that $\mathbf{A}^t C \mathbf{A} = C$. If this is so, all the $\mathbf{A}^n(R)$ must lie on the conic $v^t C v = R^t C R$, since

$$\begin{aligned} (\mathbf{A}^n(R))^t C \mathbf{A}^n(R) &= R^t(\mathbf{A}^n)^t C \mathbf{A}^n R = R^t(\mathbf{A}^t)^n C \mathbf{A}^n R \\ &= R^t(\mathbf{A}^t)^{n-1} C \mathbf{A}^{n-1} R = \cdots \\ &= R^t C R \end{aligned}$$

The matrix C, in terms of homogenous coordinates, is

$$\begin{pmatrix} -t-1 & 0 & 0 \\ 0 & t-1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

The conic that C generates is $(t-1)y^2 + 2y - (t+1)x^2 = 0$. The reader can verify that $\mathbf{A}^t C \mathbf{A} = C$.

Now if all the $\mathbf{A}^n(R)$ are different, we have generated an infinite number of point-tangents that lie on $S(P,Q)$ and which also lie on a conic. If $S(P,Q)$ is not the conic C and passes through all these point-tangents, it must have some inflection points. But inflection points are preserved under affine transformations, so we can map two tangents on $S(P,Q)$ and around these inflection points onto P and Q, which increases the number of inflection points between P and Q; so the contour is not affine commutative. Hence, if the $\mathbf{A}^n(R)$ are all different, the contour is the conic C.

If the $\mathbf{A}^n(R)$ are not all different then $\mathbf{A}$ is a root of unity (i.e., there is an n such that $\mathbf{A}^n = \mathbf{I}$, the identity matrix). In this case there will be only a finite number of points. We can use a special argument to take care of this case: map any other two tangents on $S(P,Q)$ onto P and Q, and the point crossing the y-axis onto $(0,0)$ as above. The new contour will either have the same tangents at P and Q as the old contour, or it will not.

If it has the same tangents, we keep choosing points until we get different tangents; if we never do, we have generated an infinite number of points that can be mapped by an affine transformation into the configuration P, Q, R; such a contour must be a conic, by an argument similar to the one above.

If the tangents are not the same, we use continuity to show that there exist points giving a value of k for which $\mathbf{A}$ is not a root of unity, since the class of matrices of the form of $\mathbf{A}$ which are roots of unity is countable, and the class of points on the contour is uncountable. Hence the new contour will be a conic, which means the original contour is a conic. This completes the proof. ∎

We have as a corollary:

4.3. Corollary. *The subjective contour through two non-degenerate linked point-tangents is a parabola or two straight lines.*

Proof. The class of subjective contours is closed with respect to affine transformations and translations by affine commutativity. Now we will show that no two elements of the class of subjective contours through linked point-tangents can intersect in two point-tangents. Suppose, to the contrary, that $S(P,Q) \neq S(U,V)$ are subjective contours and that they intersect in two point-tangents X and Y. We cannot have $S(X,Y) = S(P,Q)$ and $S(X,Y) = S(U,V)$, so without loss of generality assume that $S(X,Y) \neq S(U,V)$. By point replacement $S(U,V) = S(U,Y) = S(X,Y)$, which is a contradiction.

Hence the class of subjective contours through two point-tangents is closed with respect to affine transformations and has the property that no two contours from the class intersect in two point-tangents. There are only two subsets of the conics that have these properties: the parabolae and pairs of lines. ∎

The subjective contours most commonly observed are curved, so that they must be parabolae. Under special conditions straight lines may be observed, however [4]. Parabolae have been suggested by Bookstein [9] for curve interpolation.

Subjective contours can be formed by point-tangents, points, or combinations of the two. So far we have only considered the case of two linked point-tangets. In the remainder of the paper, we consider subjective contours through other kinds of data.

V. SUBJECTIVE CONTOURS THROUGH OTHER DATA

We will now complete our theory of the shape of subjective contours by considering subjective contours through non-linked point-tangents, multiple point-tangents, and points. Because of space limitations we will consider each case only briefly. We will note statements for which we have proofs, and statements which are only conjectures.

First, we consider non-linked point-tangents. If the point-tangents are parallel, the subjective contour is provably two parallel lines passing through the point-tangents. If the point-tangents are not linked but not parallel the subjective contour is provably a "double parabola," which consists of two parabola with the same shape, rotated so that they are tangent to each other and to the two point-tangents.

Next we consider multiple point-tangents. There seem to be two approaches, though we have not proved there are only two. The subjective contour can be constructed locally, by fitting subjective contours through successive pairs of point-tangents, or it can be constructed globally by fitting conics through sucessive pairs of point-tangents, and requiring curvature continuity across the point-tangents. The second approach has the advantage of producing circles as subjective contours in some situations, a condition favored by Knuth [2]. However, it is more difficult to implement. Both of these approaches provably are consistent.

Finally, we consider subjective contours through points. We conjecture that there is only one approach. Parabolae can be passed through every pair of points, and the curvature of the parabolae can be required to be continuous across the points. At the ends of the contour, we fit the parabola through the last three points on the contour, and require curvature continuity where it joins the rest of the contour. This approach is provably consistent only with the first approach to multiple point-tangents. It is provably affine commutative and allows point replacement, and we conjecture that it is extensible on the basis of demonstrations like the one in Figure 2.

All of the above applies only to curved subjective contours. The theory for straight subjective contours (consisting of intersecting straight lines) is much easier to work out. We can prove that all that is necessary is to extend point-tangents until they intersect, and to join points with straight lines. But this kind of contour-fitting is uninteresting.

We conjecture from the above discussion that curved subjective contours always consist of parabolae, with single or double parabolae fit between point-tangents, and curvature-continuous parabolae fit through points.

ACKNOWLEDGMENTS

This work benefitted from discussions with Takeo Kanade and Geoff Hinton.

REFERENCES

[1] Kanizsa, G., "Subjective contours," *Scientific American*, **234**, (4), 48-52. 1976.

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

[3] Brady, M. and W. E. L. Grimson, "Shape encoding and subjective contours," *Proceedings of the First Annual National Conference on Artificial Intelligence*, The American Association for Artificial Intelligence, Stanford, CA, 15-17. 1979.

[4] Ullman, S., "Filling-in-the-gaps: the shape of subjective contours and a model for their completion," *Biological Cybernetics*, **1**, (6). 1976.

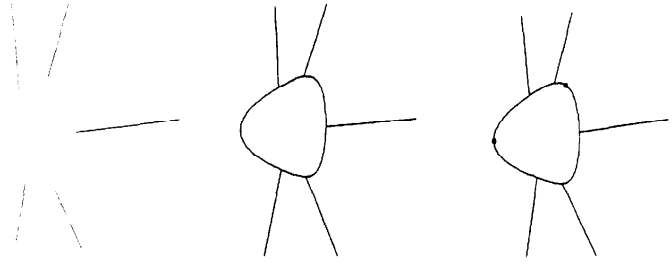
[5] Brady, M. and W. E. L. Grimson, "The perception of subjective surfaces," A. I. Memo 666, Artificial Intelligence Laboratory, Mass. Inst. Tech., Cambridge, MA. 1981.

[6] Rutkowski, W. S., "Shape completion," *Computer Graphics and Image Processing*, **9**, 89-101. 1979.

[7] Stevens, K. A., "The visual interpretation of surface contours," *Artificial Intelligence*, **17**, 47-73. 1981.

[8] Newman, W. M., and R. F. Sproull, *Principles of interactive computer graphics*, Second edition, McGraw-Hill. 1979.

[9] Bookstein, L., "Closing gaps, and gaps with a stepping stone, by means of parabolas," *Computer Graphics and Image Processing*, **10**, 372-374. 1979.



(a) (b) (c)

Figure 2. (a) Original data. (b) Interpolated contour. (c) Contour with added points (black rectangles).