IEEE COMPUTER SOCIETY
PRESS REPRINT

THE DYNAMIC ANALYSIS OF APPARENT CONTOURS

Roberto Cipolla
Andrew Blake

Reprinted from PROCEEDINGS OF THE THIRD INTERNATIONAL CONFERENCE ON COMPUTER VISION, Osaka, Japan, December 4-7, 1990

IEEE Computer Society
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1264

Washington, DC • Los Alamitos • Brussels • Tokyo

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.
IEEE

IEEE COMPUTER SOCIETY

The Dynamic Analysis of Apparent Contours

Roberto Cipolla Andrew Blake

University of Oxford
Department of Engineering Science
Parks Road
Oxford OX1 3PJ, UK.

Abstract

Robots perambulating and manipulating unmodelled environments need robust geometric cues to recover scene structure. It is furthermore attractive to capitalise on structural information inherent in the evolution of the image under robot motion. However, especially in artificial environments, surface texture may be sparse, and silhouettes or apparent contours are the dominant image features. We develop previous theories of the analysis of deformation of apparent contours under viewer motion. First, earlier results showing how surface curvature can be inferred from *acceleration* of image features are generalised for arbitrary viewer motion and perspective projection. Second, we show that *relative* image acceleration, based on parallax measurements, is robust to uncertainties in robot motion. Thirdly, our theory has been implemented and extensively tested in a realtime (15 frames per second) tracking system based on deformable contours (snakes). We show that focusing attention by means of snakes enables rapid, robust computation of surface curvature, including discrimination of extremal and occluding contours.

1 Introduction

The deformation of an apparent contour (the silhouette of a smooth surface or the image of the extremal boundary) under viewer-motion is a potentially rich source of geometric information for navigation, motion-planning and object-recognition. Barrow and Tenenbaum [1] pointed out that surface orientation along an extremal boundary can be computed directly from image data. Koenderink [15] related the curvature of an apparent contour to the intrinsic curvature of the surface (Gaussian curvature); the sign of Gaussian curvature is equal to the sign of the curvature of the contour. Convexities, concavities and inflections of an apparent contour indicate, respectively, convex, hyperbolic and parabolic surface points. Giblin and Weiss [9] have extended this by adding viewer motions to obtain quantitative estimates of surface curvature. For orthographic projection they show that a surface can be reconstructed from the envelope of all its tangent planes, which in turn are computed directly from the family of silhouettes of the surface, obtained under planar motion of the viewer. In Blake and Cipolla (1989) [3] this was extended to the general case of arbitrary non-planar, curvilinear camera motion under perspective projection.

In this paper we summarise further developments of the theory and describe the implementation and results with a camera mounted on a moving robot arm. We describe a simple, computationally efficient method for accurately extracting image curves from real images and tracking their temporal evolution. This is an extension of tracking with *Snakes* [14] - energy minimising splines guided by image forces - in which we avoid computing the internal energies by representing sections of curves by cubic B-splines and can achieve realtime processing (15 frames per second) by windowing [13]. Experiments show that with adequate viewer motion calibration it is possible to obtain 3D shape measurements of useful accuracy.

A consequence of the theory, representing an important step towards qualitative vision, concerns the robustness of measurements of curvature based on *motion parallax* at two nearby points. Intuitively it is relatively difficult to judge, moving around a smooth, featureless object, whether its silhouette is extremal or not — whether curvature along the contour is bounded or not. This judgment is much easier to make for objects which have at least a few surface features. Under small viewer-motions, features are "sucked" over the extremal boundary, at a rate which depends on surface curvature. Our theoretical findings exactly reflect the intuition that the "sucking" effect is a reliable indicator of relative curvature, regardless of the exact details of the viewer's motion. Relative measurements of curvature across two adjacent points are entirely immune to uncertainties in the viewer's rotational velocity. This is somewhat related to earlier results showing that relative measurements of this kind are important for depth measurement from optic flow [16, 19, 21] and for curvature measurements from stereoscopically viewed highlights [2]. Furthermore, they are relatively immune to uncertainties in translational motion in that, unlike single-point measurements, they are independent of the viewer's acceleration. Only dependence on velocity remains. Experiments show that this theoretical prediction is borne out in practice. Surface curvature estimated from parallax measurements prove to be more than an order of magnitude less sensitive than single-point measurements to errors in viewer-motion calibration.

As an illustration of their power, we show how these motion analysis techniques can achieve something which has so far eluded photometric analysis: namely reliable discrimination between fixed surface features and points on extremal boundaries and the reconstruction of surfaces in the vicinity of their extremal boundaries.

CH2934-8/90/0000/0616$01.00 © 1990 IEEE

2 Theoretical framework

It is now well established that static views of extremal boundaries are rich sources of surface geometry [1, 5, 9, 15]. The physical constraints of *tangency* (all rays at an extremal boundary are in the surface's tangent plane) and *conjugacy* (the extremal boundary is not in general orthogonal to the ray direction but *conjugate* to it) allow the recovery of surface orientation and the sign of Gaussian curvature directly from the image of the extremal boundary, the apparent contour [3].

Moreover, each vantage point, generates a new extremal boundary making it an ideal cue in the active exploration of 3D geometry. We outline below how the deformation of apparent contours under known viewer-motion can be used to recover the position, orientation and full surface curvature (3D shape) of visible surfaces in the vicinity of their extremal boundaries.

2.1 Surface Geometry

Consider a point P on the extremal boundary of a smooth, curved surface which we represent locally by a vector valued function $\mathbf{r}(s,t)$ with parameters s and t. The parametric representation can be considered as covering the surface with 2 families of curves: $\mathbf{r}(s,t_0)$, and $\mathbf{r}(s_0,t)$ where s_0, t_0 are fixed for a given curve in the family. A one-parameter family of views is indexed by the time parameter t and s,t are defined so that the s-parameter curve, $\mathbf{r}(s,t_0)$, is an extremal boundary for a particular view t_0 (figure 1).

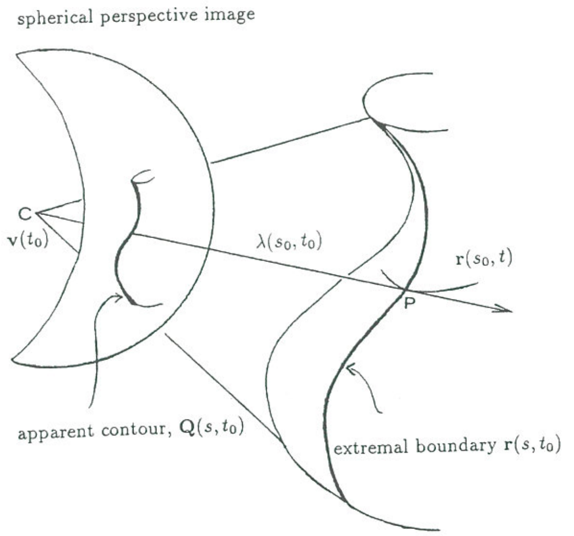

Figure 1. Surface and Viewing Geometry.
P lies on a smooth surface which is parameterised locally as $\mathbf{r}(s,t)$. *For a given vantage point,* $\mathbf{v}(t_0)$, *the family of rays emanating from the viewer's optical centre (C) that touch the surface defines an s-parameter curve* $\mathbf{r}(s,t_0)$ *– the extremal boundary from vantage point* t_0. *The spherical perspective projection of this extremal boundary – the apparent contour,* $\mathbf{Q}(s,t_0)$ *– determines the direction of rays which graze the surface. The distance along the ray, CP, is* $\lambda(s_0,t_0)$.

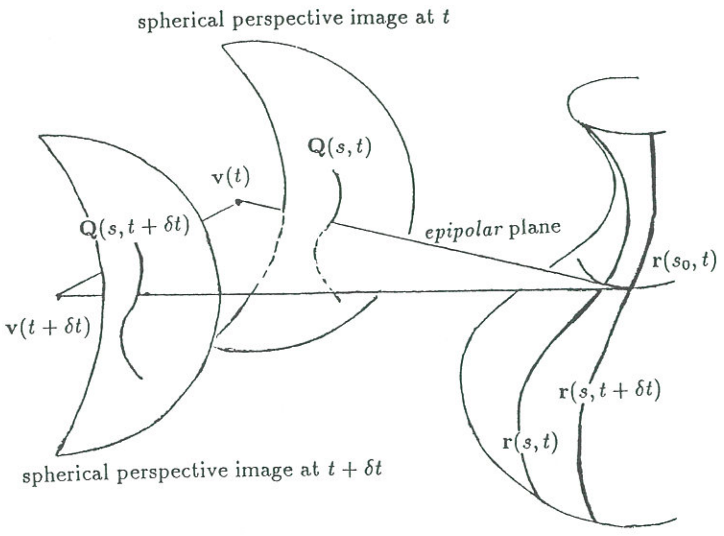

Figure 2. Epipolar parameterisation
A moving observer at position $\mathbf{v}(t)$ *sees a family of extremal boundaries* $\mathbf{r}(s,t)$ *indexed by the time parameter* t. *Their spherical perspective projections are represented by a 2 parameter family of apparent contours* $\mathbf{Q}(s,t)$. *For the epipolar parameterisation t-parameter curves (*$\mathbf{r}(s_0,t)$ *and* $\mathbf{Q}(s_0,t)$*) are defined by chosing the correspondence between successive contours to be in an epipolar plane which is determined by the translational velocity and the direction of the ray.*

A t-parameter curve $\mathbf{r}(s_0,t)$ can be thought of as the 3D locus of points grazed by a light-ray from the viewer, under viewer-motion. Such a locus is not uniquely defined. The correspondence, as the viewer moves, between "successive" (in an infinitesimal sense) extremal boundaries is not unique. Hence there is considerable freedom to choose a spatio-temporal parameterisation of the surface, $\mathbf{r}(s,t)$. A natural choice of parametrisation, it has been shown [3], is the *epipolar parameterisation* in which points on successive extremal boundaries are matched in the epipolar plane defined by the ray direction and the instantaneous viewer translational velocity (figure 2). The advantage of the parametrisation is clear below, when it leads to a simplified treatment of surface curvature and a unified treatment of the projection of rigid space curves and extremal boundaries.

Surface curvature (3D shape) can be expressed in terms of the coefficients of the first and second fundamental forms $\mathbf{I}, \mathbf{II}$ [8]. For the *epipolar* parametrisation, these are:

$$\mathbf{I} = \begin{bmatrix} 1 & cos\theta \\ cos\theta & 1 \end{bmatrix} \qquad (1)$$

$$\mathbf{II} = \begin{bmatrix} \kappa^t & 0 \\ 0 & \kappa^s \end{bmatrix}, \qquad (2)$$

where θ is the angle between the ray direction and the extremal boundary; κ^t is the normal curvature of the t-parameter curve $\mathbf{r}(s_0,t)$) and κ^s is the normal curvature of the extremal boundary $\mathbf{r}(s,t_0)$ at P. Equivalently κ^t is the curvature of the normal section at P in the direction of the ray. Note that $\mathbf{II}$ is diagonal, a result of choosing, in the epipolar parametrisation, basis directions that are *conjugate* [7].