1

Introduction

We have tried to make our book self-contained. The underlying differential geometry and singularity theory is explained with minimal prerequisites in Chapter 2. We hope that this chapter will prove of value to anyone who wishes to apply differential geometry to vision problems. To follow the main thrust of this chapter the reader only needs a working knowledge of calculus and linear algebra. Most of the material is quite well known, but the slant given here is towards applications, and we have tried to illustrate the material with many examples and figures. In particular, we study curvature of surfaces and special curves on surfaces such as parabolic and flecnodal curves. We make much use of the idea of contact, between surfaces and lines or planes. It is this idea which links classical differential geometry with modern singularity theory, where geometrical properties are studied by means of functions or mappings which in turn measure contact.

In Chapter 3 we introduce the main character in our story, the apparent contour. Apparent contours are the outlines or profiles of curved surfaces. An example is shown in Figure 1.1. We describe apparent contours under orthographic projection and for perspective projection, and we describe in detail the singularities which a single apparent contour can be expected to have. We also obtain geometrical information about surfaces from a single apparent contour, though this is necessarily limited. Initially results are stated, but we give three different approaches to proofs, 'Monge–Taylor proofs', which rely on special coordinate systems which are very powerful for proving results about surfaces; 'vector proofs', which are coordinate-free but require more experience to use effectively; and 'pure geometric proofs', which are more like thought experiments but can sometimes yield the greatest intuition.

An excellent modern reference for applications of projective geometry to computer vision is O.D. Faugeras' *Three-Dimensional Computer Vision*. On the differential geometry side, another book from which we, and others, have drawn inspiration is J.J. Koenderink's now classic *Solid Shape*, published in 1990. Koenderink's book is replete with geometric proofs and statements, but sometimes lacks mathematical detail. We have tried to supply some of this detail in Chapters 2 and 3 of our book.

In Chapter 4 we introduce dynamic contours. That is, we progress from a single view of a surface to multiple views, from which we can expect to derive much more information. In fact, in principle, a complete reconstruction of a surface is possible from a family of apparent contours obtained by circumnavigating a surface. (Unfortunately in practice some parts of a surface may be occluded by other parts, and in addition apparent contours may be hard to track.) The 'in principle' reconstruction was first established for orthographic projection in Giblin and Weiss (1987), and this was generalized and placed in a better mathematical framework by Cipolla and Blake (1990 and 1992).

We describe the dynamic analysis for orthographic and perspective projection, and introduce the important idea of an epipolar parametrization. We also give a brief introduction to circular motion, which will play a major role in the last chapters of the book. The epipolar parametrization breaks down in certain circumstances, one of which is the 'epipolar tangency' situation. This is bad news for reconstruction but, surprisingly, very good news for determining motion. In fact the so-called frontier points which arise from epipolar tangency are instrumental in giving us information about the motion of the observer, something we exploit in Chapter 6. Other breakdowns of the epipolar parametrization are caused by degeneracies of the apparent contour – the 'visual events' which we observe when moving our viewpoint – and we list the possible cases and explain their geometrical significance.

In Chapter 5 we bring the mathematical techniques to life and describe the implementation of algorithms to reconstruct a surface from the image sequence of outlines. Details of every stage in the reconstruction, from raw pixel intensities to a stable description of the three-dimensional surface, are given. These include the calibration of cameras, localization and tracking of outlines, epipolar geometry and stereo reconstruction.

In Chapter 6 we address the more difficult problem of recovering the observer's motion from the apparent contours in different views. The recovery of the three-dimensional configuration of points and the motion compatible with their views (known as structure from motion) has been an active area of research in computer vision over the last two decades and a large number of algorithms and working systems already exist. We review the key results in the literature, in many cases providing simple geometric and algebraic proofs. Finally we show how the motion of the viewer can be computed

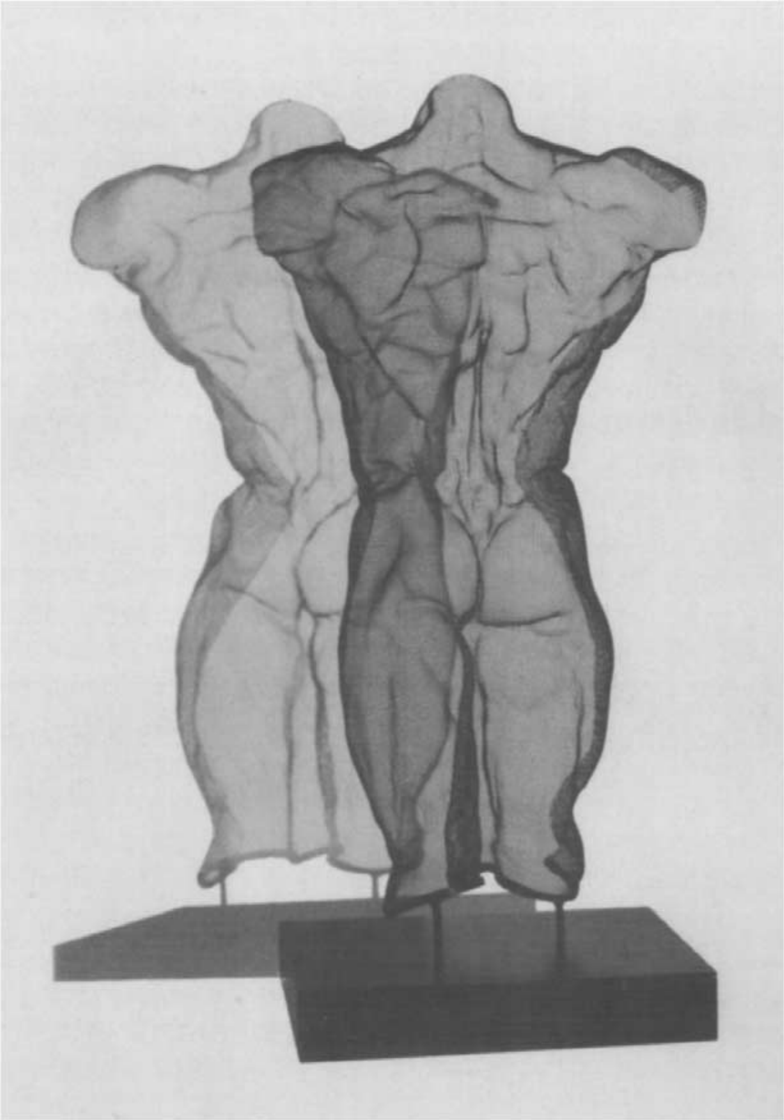

Fig. 1.1 Two views of a semi-transparent surface (sculpture) by David Begbie. For curved surfaces, the dominant image feature is the apparent contour or outline. This is the projection of the locus of points on the surface which separate visible and occluded parts. The apparent contours are rich sources of geometric information. In particular, their deformation under viewer motion can be used to recover the geometry of the visible surface. The geometry of the viewpoints can also be inferred.

from apparent contours instead of points, by using properties of the frontier. The significance of frontier points seems to have been noticed first by J. Rieger (1986), and they were then applied to circular motion and orthographic projection in (Giblin et al. 1994), where it is proved that recovery of motion is essentially unique in this simple case. The extension to general motion and perspective projection was presented by (Cipolla et al. 1995) and (Aström et al. 1996 and 1999), where an iterative algorithm gives good results in many cases. We present the latest techniques for estimating the camera motion. A particularly simple and reliable method is presented for recovering the motion of objects on turntables, known as circular motion. This exploits symmetry of the envelope of apparent contours. This has been used to acquire three-dimensional models of arbitrary objects from an uncalibrated camera.

2

Differential Geometry of Curves and Surfaces

In this chapter we aim to introduce the reader to all the differential geometry of curves and surfaces in 3-space needed for a full understanding of the remainder of the book, and of the literature on apparent contours and their applications in computer vision. Our aim is to give all the useful formulae and to make the concepts and results clear. We give several methods of proof, and in some cases apply all these methods to the same problem, to give a flavour of the strengths and weaknesses of the different techniques available. We begin with 'first-order properties', that is properties which depend only on first derivatives, and work up to those depending on second or higher derivatives.

2.1 Curves and their tangents

A curve in 3-dimensional space is nearly always represented by a parametrization

$$\mathbf{r} : I \to \mathbf{R}^3,$$

where I is some interval $a < t < b$ of real numbers (a could be $-\infty$ or b could be ∞, or both). Also I could be a circle for parametrizing a closed curve, in which case the components of $\mathbf{r}$ will be functions of $\sin t$ and $\cos t$. Writing $\mathbf{r}(t) = (X(t), Y(t), Z(t))$ the curve is *regular* provided the *velocity vector* $\mathbf{r}'(t) = (X'(t), Y'(t), Z'(t))$ is never the zero vector. (Here the prime $'$ stands for $\frac{d}{dt}$. We might also use a suffix: $\mathbf{r}_t = \mathbf{r}'$, although we normally reserve suffixes for partial derivatives where there is more than one variable.) The *unit tangent vector* $\mathbf{T}(t)$ is the unit vector in the direction of the velocity, namely

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{||\mathbf{r}'(t)||}.$$

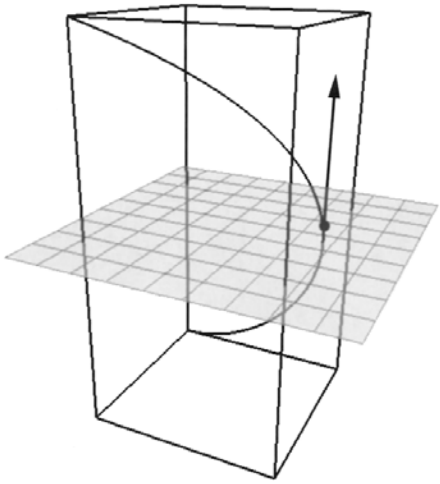

Fig. 2.1. A normal plane and tangent vector of a space curve.

The denominator, $||\mathbf{r}'(t)||$, is called the *speed* of the curve; *unit speed* means that $||\mathbf{r}'(t)|| = 1$ for all t. The curve has a *normal plane* at any point $\mathbf{r}(t_0)$, namely the plane through this point perpendicular to the tangent vector. See Figure 2.1. This plane has equation $(\mathbf{x} - \mathbf{r}(t_0)) \cdot \mathbf{r}'(t_0) = 0$.

The *arclength* $s = l(t)$ of the curve between $\mathbf{r}(t_0)$ and $\mathbf{r}(t)$ is given by

$$s = l(t) = \int_{t_0}^{t} ||\mathbf{r}'(t)||\, dt, \tag{2.1}$$

that is, the integral of the speed. We deduce from (2.1) that $\frac{ds}{dt} = ||\mathbf{r}'(t)||$. In particular if the parameter t is s itself, then $||\mathbf{r}'(s)|| = 1$: using arclength as the parameter, the curve is automatically of unit speed. In order that changing from t to s is a valid change of parameter we need to know that $\frac{ds}{dt}$ is never zero. Thus every *regular* curve can be re-parametrized to be unit speed. This is often useful when doing calculations. An example of a non-regular curve is given by $\mathbf{r}(t) = (t^2, t^3, 0)$, which is a curve in the x, y-plane with a cusp at the origin. The speed is zero precisely for $t = 0$.

Example 2.1.1 Helix

Let $\mathbf{r}(t) = (\cos t, \sin t, t)$, which has $\frac{ds}{dt} = ||\mathbf{r}'(t)|| = \sqrt{2}$ for all t. Thus $s = \sqrt{2}t +$ constant and, making $s = 0$ when $t = 0$, the reparametrized curve

$$\mathbf{R}(s) = \left(\cos\frac{s}{\sqrt{2}}, \sin\frac{s}{\sqrt{2}}, \frac{s}{\sqrt{2}}\right)$$

is unit speed. □

2.2 Surfaces: the parametric form

With a *surface* the matter is quite different. There is no natural concept of 'unit speed'. The geometry of surfaces is also much richer than that of curves, and it is well worth having several representations of a surface available for different purposes. We shall describe three of these.

We take a plane—the *parameter plane*—with coordinates (u, v) and an open and connected region U of this plane. Then we can parametrize a surface by

$$\mathbf{r}(u, v) = (X(u, v), Y(u, v), Z(u, v)),$$

where X, Y, Z are functions with sufficiently many continuous derivatives (at least two and often infinitely many). The condition corresponding to the 'regularity' of a space curve is that, for every $(u, v) \in U$, the *Jacobian matrix*

$$\begin{pmatrix} X_u & X_v \\ Y_u & Y_v \\ Z_u & Z_v \end{pmatrix}$$

(where suffixes stand for partial derivatives) should have rank 2. The content of this is that the rank should not drop *below* 2. If it does drop, the two columns $\mathbf{r}_u$ and $\mathbf{r}_v$ are parallel (or zero) as vectors in $\mathbf{R}^3$. Note that

$$\mathbf{r}_u,\ \mathbf{r}_v \text{ are parallel or zero } \Leftrightarrow \mathbf{r}_u \wedge \mathbf{r}_v = \mathbf{0}.$$

In fact $\mathbf{r}_u$, evaluated at (u, v_0), is the tangent vector to the space curve $\mathbf{r}(u, v_0)$ where v has a fixed value v_0, and similarly $\mathbf{r}_v$ is the tangent vector to the space curve $\mathbf{r}(u_0, v)$. If these two vectors are parallel then the perpendicular coordinate curves $u = u_0, v = v_0$ are transformed into tangential curves by $\mathbf{r}$, which is not allowable.

When the Jacobian matrix always has rank 2, i.e. $\mathbf{r}_u \wedge \mathbf{r}_v$ is never the zero vector, we say that the map $\mathbf{r}$ is an *immersion*, and that $\mathbf{r}$ defines an *immersed surface*, namely $M = \mathbf{r}(U)$. The surface M has a *tangent plane* at $\mathbf{r}(u, v)$ for every parameter point (u, v), namely the plane through $\mathbf{r}(u, v)$ spanned by the vectors $\mathbf{r}_u$ and $\mathbf{r}_v$. This is a genuine 2-dimensional plane by the immersion condition. So an immersed surface can be thought of as one where, corresponding to every parameter point (u, v), there is a definite tangent plane at $\mathbf{r}(u, v)$.

All surfaces will be assumed immersed unless the contrary is stated.

Note that $\mathbf{r}_u$ and $\mathbf{r}_v$ are not perpendicular in general. The normal to the tangent plane is also called the *normal* to the surface and is in the direction $\mathbf{r}_u \wedge \mathbf{r}_v$. See Figure 2.2. The *unit surface normal* $\mathbf{n}$ determined by the ordering

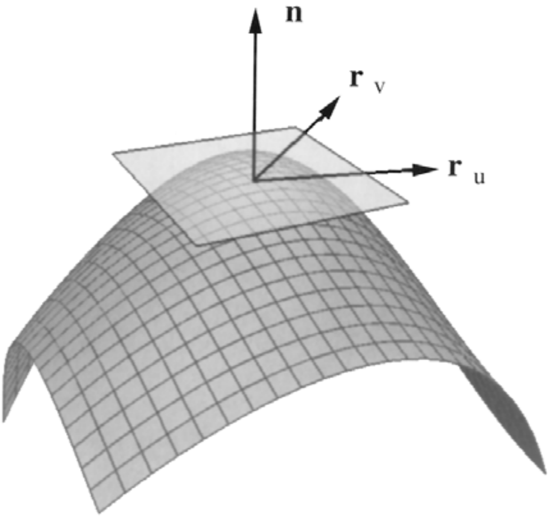

Fig. 2.2 Surface normal and tangent vectors $\mathbf{r}_u, \mathbf{r}_v$. Note that the lines drawn are not u, v coordinate curves; they just give a shape to the surface.

u, v of the parameters is

$$\mathbf{n} = \frac{\mathbf{r}_u \wedge \mathbf{r}_v}{||\mathbf{r}_u \wedge \mathbf{r}_v||}.$$

Note that interchanging u and v takes $\mathbf{n}$ into $-\mathbf{n}$, and that the normal depends only on first derivatives of the parametrization.

Sometimes it is convenient to use a slightly different parameter space, for example where one, or both, of the parameters is naturally taken to be an angle. (Formally we would say the parameter space was the product of a circle and a line, or of two circles.)

Example 2.2.1 Sphere

A unit sphere, centre $(0,0,0)$, minus the north pole $(0,0,1)$ can be parametrized by *stereographic projection* as

$$\mathbf{r}(x, y) = \left(\frac{4x}{x^2+y^2+4}, \frac{4y}{x^2+y^2+4}, \frac{x^2+y^2-4}{x^2+y^2+4} \right)$$

where (x, y) ranges over the whole plane. Here, the point $(x, y, -1)$ in the plane $z = -1$ is joined to the point $(0,0,1)$ and this segment meets the sphere again at $\lambda(x, y, -1) + (1-\lambda)(0,0,1)$ where $\lambda = 4/(x^2+y^2+4)$. See Figure 2.3. It can be shown that $\mathbf{r}_x \wedge \mathbf{r}_y$ is never zero, so this $\mathbf{r}$ is an immersion.

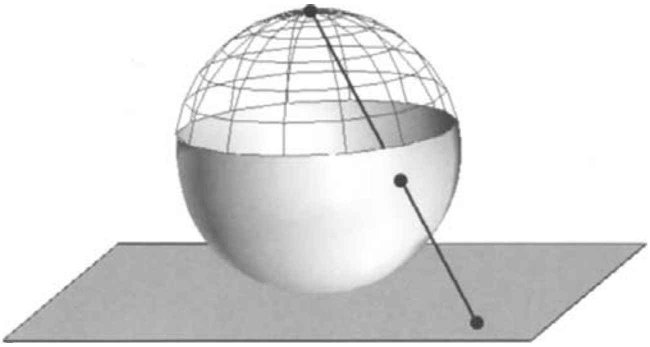

Fig. 2.3 Stereographic parametrization of a sphere minus the north pole. Note that the sphere is cut away simply to show the line from the north pole penetrating the sphere; the whole sphere minus the north pole is covered.

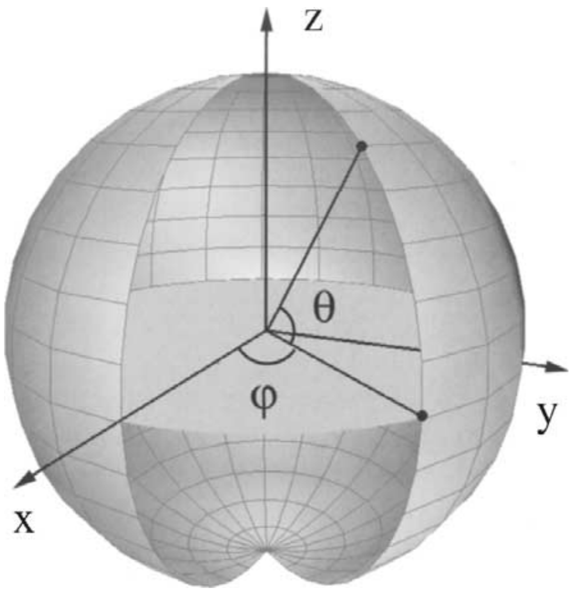

Fig. 2.4 Cut away picture of θ (latitude) and ϕ (longitude) coordinates on the sphere minus north and south poles.

A sphere, radius ρ, minus the north and south poles, can be parametrized by spherical polar coordinates as $\mathbf{r}(\theta, \phi) = (\rho\cos\theta\cos\phi, \rho\cos\theta\sin\phi, \rho\sin\theta)$ where $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$, $0 \leq \phi < 2\pi$. (Here, ϕ is regarded as an angle parameter which covers a whole circle.) See Figure 2.4. The tangent vectors and unit normal are

$$\mathbf{r}_\theta = \begin{bmatrix} -\rho\sin\theta\cos\phi \\ -\rho\sin\theta\sin\phi \\ \rho\cos\theta \end{bmatrix} \qquad \mathbf{r}_\phi = \begin{bmatrix} -\rho\cos\theta\sin\phi \\ \rho\cos\theta\cos\phi \\ 0 \end{bmatrix} \qquad \mathbf{n} = \begin{bmatrix} -\cos\theta\cos\phi \\ -\cos\theta\sin\phi \\ -\sin\theta \end{bmatrix}.$$

The normal is (of course!) parallel to $\mathbf{r}$. Note that if $\theta = \pm\frac{\pi}{2}$ then $\mathbf{r}_\phi = \mathbf{0}$; this is why we exclude the north and south poles from the parametrization. □

Example 2.2.2 Cylinder

A cylinder, radius ρ, with axis along the z-axis, can be parametrized by cylindrical polar coordinates as $\mathbf{r}(\phi, z) = (\rho\cos\phi, \rho\sin\phi, z)$. (Again, ϕ covers a whole circle here.) The tangent vectors and unit normal are

$$\mathbf{r}_\phi = \begin{bmatrix} -\rho\sin\phi \\ \rho\cos\phi \\ 0 \end{bmatrix} \qquad \mathbf{r}_z = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad \mathbf{n} = \begin{bmatrix} \cos\phi \\ \sin\phi \\ 0 \end{bmatrix}.$$

The normal is orthogonal to the cylinder's axis. □

Example 2.2.3 A non-surface

Let $\mathbf{r}(u, v) = (u + 2v, 2u + 4v, 3u + 6v)$. Although $\mathbf{r}$ is parametrized by two parameters, the tangent vectors

$$\mathbf{r}_u = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{r}_v = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$$

are parallel, so a surface normal is undefined. In fact this $\mathbf{r}$ gives a curve—indeed a straight line—not a surface. □

Example 2.2.4 Almost an immersed surface: the crosscap

A less drastic failure—which we include here simply to show what the immersion definition allows—is $\mathbf{r}(u, v) = (u, v^2, uv)$. Here, $\mathbf{r}_u = (1, 0, v)^\top, \mathbf{r}_v = (0, 2v, u)^\top$ and these are parallel (vector product zero) if and only if $u = v = 0$. So excluding the origin $(0, 0)$ from the parameter plane, $\mathbf{r}$ gives an immersed surface. See Figure 2.5. The surface is called a 'crosscap' or 'Whitney umbrella'[1] and has a line of self-intersection corresponding to $u = 0$: we have $\mathbf{r}(0, v) = \mathbf{r}(0, -v)$ for all v. Thus at points $(0, v^2, 0)$ $(v \neq 0)$ there are

[1] The handle of the umbrella is actually the negative y-axis, which is not present in this parametrized form but is present in the 'equation' form $x^2y = z^2$.