Figure 6.1. Cinderella trying on the slipper in Gustave Doré's engraving.

6.1 Moments of joy, moments of sorrow

Every rigid transformation in $\mathbb{R}^3$ can be described by six parameters: three rotation angles $\theta = (\theta^1, \theta^2, \theta^3)^{\mathrm{T}}$ about the x, y, and z axes, respectively, and three translation coordinates $t = (t^1, t^2, t^3)^{\mathrm{T}}$. Such a transformation repositions a vector x in $\mathbb{R}^3$ to

$$x' = Rx + t \;=\; R_1 R_2 R_3 x + t,$$

where

$$R_1 \;=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta^1 & \sin\theta^1 \\ 0 & -\sin\theta^1 & \cos\theta^1 \end{pmatrix}, R_2 \;=\; \begin{pmatrix} \cos\theta^2 & 0 & \sin\theta^2 \\ 0 & 1 & 0 \\ -\sin\theta^2 & & \cos\theta^2 \end{pmatrix},$$

and

$$R_3 = \begin{pmatrix} \cos\theta^3 & \sin\theta^3 & 0 \\ -\sin\theta^3 & \cos\theta^3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

are rotation matrices.[2]

A straightforward approach for getting rid of rigid isometries is to find a Euclidean transformation that brings a surface X to some "canonical" placement in $\mathbb{R}^3$. For example, if we could identify a landmark point s_0 on X, translating the surface by $t = -s_0$ would always bring that point to the origin, resolving the ambiguity in surface position. However, finding landmark points requires additional information about the surface, which is not always available.

Nevertheless, there exist several points that can be found for every three-dimensional surface. One of such points is the *extrinsic centroid* (the terms *center of mass* and *center of gravity* are often used as synonyms),

$$x_0 = \int_X x dx, \tag{6.1}$$

which is essentially the "average location" of X (note that unlike its intrinsic counterpart we have encountered in Chapter 3, the extrinsic centroid does not necessarily belong to X). Clearly, translating the surface in such a way that x_0 coincides with the origin resolves the translation ambiguity.

Next, we have to resolve the remaining three degrees of freedom due to rotation. This can be done by finding a direction in which the surface has maximum extent, and aligning it, say, with the the e_1 axis (Figure 6.2, left). Because a direction is described by a unit vector in $\mathbb{R}^3$, this step resolves only two of the three degrees of freedom. The remaining degree of freedom is due to the rotation ambiguity about the e_1 axis. However, we can apply the same idea again by rotating the surface such that the projection on the e_2e_3 plane, which can be illustrated as the footprint of the shadow cast by the surface (Figure 6.2, right), has the maximum extent in the direction of the e_2 axis.

Formally, the first direction we are looking for can be defined as the one that maximizes the *variance* of the projection of X onto it,

$$d_1 = \arg \max_{d_1 : \|d_1\|_2 = 1} \int_X (d_1^{\mathrm{T}} x)^2 dx,$$

where we assume that the surface has already been translated so that $x_0 = 0$. Observe that the integrand $(d_1^{\mathrm{T}} x)^2$ can be written as $d_1^{\mathrm{T}} x x^{\mathrm{T}} d_1$. Because d_1 does not participate in the integration, we can write

$$d_1 = \arg \max_{d_1 : \|d_1\|_2 = 1} d_1^{\mathrm{T}} \left(\int_X x x^{\mathrm{T}} dx \right) d_1 \;=\; \arg \max_{d_1 : \|d_1\|_2 = 1} d_1^{\mathrm{T}} \Sigma_X d_1.$$

Σ, is a 3×3 matrix, whose elements

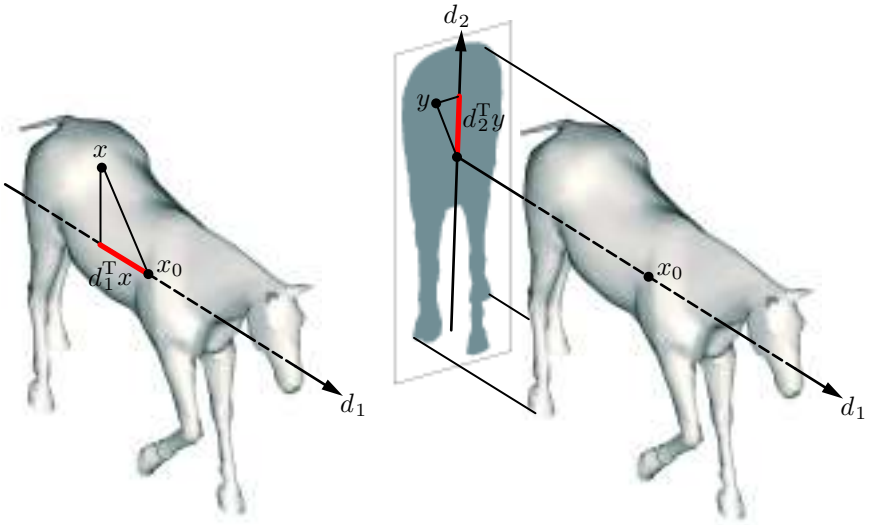

Figure 6.2. The first principal direction d_1 of the surface maximizes the variance of the projection of X onto it (left). Then, the surface is projected onto the plane orthogonal to d_1 (right, grayed) and the second principal direction d_2 is chosen as the maximum variance direction in that plane.

$$\sigma_{ij} = \int_X x^i x^j dx \tag{6.2}$$

are usually referred to as the *second-order geometric moments* of the surface,[3] and the direction d_1 maximizing the projection variance is called the first *principal direction*. Observe that the first principal direction, which has to maximize $d_1^{\mathrm{T}} \Sigma d_1$, is nothing but the first eigenvector of Σ corresponding with its maximum eigenvalue. In order to find the second principal direction, we have to project the surface onto the plane orthogonal to d_1 and find the vector d_2 in that plane, which maximizes the variance of the projection. Obviously, d_2 corresponds with the second largest eigenvector of Σ.

Because the matrix Σ is symmetric, it admits unitary diagonalization, that is, $\Sigma = U^{\mathrm{T}} \Lambda U$, where Λ is a diagonal matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ of Σ along the diagonal, and U is a unitary matrix whose columns are the corresponding eigenvectors. We leave as an exercise (Problem 6.1) the proof of the fact that U^{T} is a rotation matrix aligning d_1 and d_2 with the e_1 and e_2 axes, respectively. Clearly, after such an alignment, the *main* second-order moments σ_{ii} coincide with λ_i, whereas the *mixed* second-order moments (that is, the off-diagonal elements σ_{12}, σ_{13} and σ_{23}) vanish.

Thus far, we have seen that the transformation $(R, t) = (U^{\mathrm{T}}, -U^{\mathrm{T}} x_0)$ resolves the ambiguity of rigid isometries and brings the surface into a "canonical" configuration in the Euclidean space (Figure 6.3). Our goal is now to

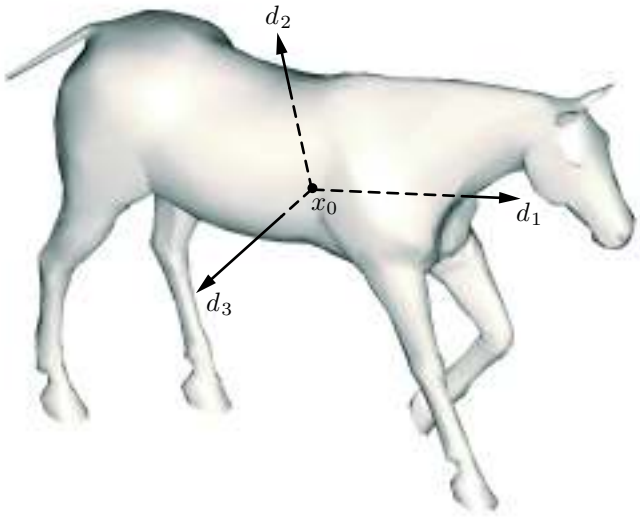

Figure 6.3. The two principal directions d_1, d_2 and a unit vector d_3 orthogonal to them define a natural coordinate system of the surface. Aligning these principal directions with the axes e_i of the standard Euclidean basis resolves the rotation ambiguity.

compare between two surfaces X and Y and quantify their similarity. We observe that the three eigenvalues λ_1, λ_2, and λ_3 of Σ provide some information about the surface extrinsic geometry. Indeed, a shape similar to a sphere is expected to have $\lambda_1 \approx \lambda_2 \approx \lambda_3$, whereas a more elongated surface should definitely have $\lambda_1 \gg \lambda_2$. In other words, the ratios $\lambda_2 : \lambda_1$ and $\lambda_3 : \lambda_1$ describe the shape *eccentricity*, and their magnitude express the shape *scale*.

We do not have to stop at the second-order moments and can define the $(p+q+r)$-th order geometric moment as

$$m_{pqr} = \int_X (x^1)^p (x^2)^q (x^3)^r dx. \tag{6.3}$$

Note that the center of gravity of the surface is a vector of its first-order moments, $x_0 = (m_{100}, m_{010}, m_{001})^{\mathrm{T}}$, whereas the elements of Σ correspond with $\sigma_{11} = m_{200}, \sigma_{22} = m_{020}, \sigma_{33} = m_{002}$ (diagonal elements), and $\sigma_{12} = m_{110}, \sigma_{13} = m_{101}, \sigma_{23} = m_{011}$ (off-diagonal elements). Higher-order moments depend on the surface position and orientation; they should be computed after performing the alignment step that eliminates the first-order and mixed second-order moments. The discretization of the integral in equation (6.3) is left as an exercise to the reader (Problem 6.3).

Intuitively, higher-order geometric moments provide us information about the surface: the more m_{pqr}'s we take, the better we can identify our object. It appears that if all moments of two surfaces coincide, the surfaces are identical. In order to understand this property, let us rewrite the (p, q, r) geometric moment of a surface as

$$m_{pqr}(f) = \int_{\mathbb{R}^3} \psi_{pqr}(x) f(x) dx \;=\; \langle \psi_{pqr}, f \rangle, \tag{6.4}$$

where $\psi_{pqr}(x) = (x^1)^p (x^2)^q (x^3)^r$, and $f : \mathbb{R}^3 \to \mathbb{R}$ is a superposition of characteristic functions, taking the value of "infinity" for $x \in X$ and zero elsewhere[4] in $\mathbb{R}^3$. Using these notations, we immediately notice that $\{m_{pqr}\}_{p,q,r=0}^{\infty}$ assume the role of the decomposition coefficients of f in the set of monomial functions $\{\psi_{pqr}\}_{p,q,r=0}^{\infty}$. Because $\{\psi_{pqr}\}$ span the space of all finite energy (more precisely, square integrable or L^2) functions on $\mathbb{R}^3$, the set of coefficients $\{m_{pqr}\}$ is unique for each surface. Indeed, if the functions f and g describing two surfaces X and Y, respectively, differ by some $h = f - g$ with non-zero energy (that is, $\int_{\mathbb{R}^3} h^2(x) dx > 0$), then there must exist some non-zero coefficients $m_{pqr}^h(h)$ such that $h = \sum m_{pqr}(h) \psi_{pqr}$. Consequently,

$$m_{pqr}(f) = \langle \psi_{pqr}, f \rangle \;=\; \langle \psi_{pqr}, g + h \rangle \;=\; m_{pqr}(g) + m_{pqr}(h) \;\neq\; m_{pqr}(f),$$

at least for some values of p, q, and r. This means that the set of all geometric moments constitutes a unique descriptor of a given surface, which is also invariant to rigid isometries if proper alignment is performed. This descriptor is also complete, meaning that, at least theoretically, one can recover[5] the surface from $\{m_{pqr}\}_{p,q,r=0}^{\infty}$.

Generally, all moments are needed to uniquely identify a surface. If we are given only a truncated set $\{m_{pqr}\}_{p,q,r=0}^{P}$ of moments up to the P-th order, there exist an infinitely large number of surfaces differing only in moments above the P-th order. However, it appears that this variety of surfaces becomes more and more similar to our surface as we increase P. In other words, even a finite set of high-order moments serves as a "fingerprint" or "signature" that identifies a sufficiently narrow class of surfaces. Ideally, we would like to be able to say that surfaces with bounded "frequencies" can be uniquely described by a finite set of moments.[6] Unfortunately, in the case of geometric moments, it is difficult to express the geometric properties of such surfaces. For this reason, geometric moments are not the best choice for measuring similarity of shapes. Other types of moments having a more clear "frequency" interpretation such as the spherical harmonics [188] or the Legendre moments [376] are usually preferred.

Using a finite set of moments, either geometric or other, we can quantify the similarity of two surfaces X and Y by applying some norm to the difference between their finite moment signatures $\{m_{pqr}(X)\}$ and $\{m_{pqr}(Y)\}$, for example,

$$d_{\mathrm{MOM}}(X,Y) = \sum_{p,q,r=0}^{P} (m_{pqr}(X) - m_{pqr}(Y))^2. \tag{6.5}$$

Said differently, d_{MOM} is a *distance function* that measures the dissimilarity between two surfaces (hereinafter, we use the term "distance" in a broad sense, not necessarily implying that d_{MOM} is a metric). Provided that X and Y are aligned prior to computing d_{MOM}, this distance function is invariant to rigid isometries. Surfaces having small distance between them are supposed to be nearly congruent (extrinsically similar), and conversely, nearly congruent surfaces result in a small d_{MOM}.

However, it is important to mention that the moment signature distance d_{MOM} has several flaws. First, recall that the continuous surfaces X and Y that we have been using freely are never available; all we have are samplings of the surfaces. It appears that the computation of moments is sensitive to the sampling, or more precisely, to sampling non-uniformity. Second, a relatively dense sampling is required in order to obtain reliable results. Third, computation of high-order geometric moments is sensitive to acquisition noise and inaccuracies due to the use of finite-precision arithmetics (see Problem 6.5). These shortcomings may limit the applicability of surface comparison methods based on moment signatures. Yet, a more serious disadvantage of d_{MOM} is that we cannot use it as a criterion of *partial similarity*.

Returning to our fairy tale example, imagine that the Prince imprudently drops the glass slipper, which breaks apart. Using moments signatures, he would never succeed in finding Cinderella, as a part of the slipper obviously has different moments than the does complete one. It is clear that the Prince needs a better distance function that still works even when the surfaces are given only partially. To his help comes a family of the so-called *iterative closest point* algorithms (*ICP* for short), first introduced by Chen and Medioni [99], and then independently by Besl and McKay [31].

6.2 Iterative closest point algorithms

The idea behind the iterative closest point algorithms is simple: given two surfaces, X and Y, find the rigid transformation (R,t), such that the transformed surface $Y' = RY + t$ is as "close" as possible to X. "Closeness" is expressed in terms of some *surface-to-surface distance* $d(RY + t, X)$. More precisely, ICP can be formulated as the minimization problem,

$$d_{\mathrm{ICP}}(X,Y) = \min_{R,t} d(RY + t, X). \tag{6.6}$$

The minimum surface-to-surface distance expresses the extrinsic similarity of X and Y. Because the minimum is searched over all Euclidean transformations, d_{ICP} is clearly invariant to rigid isometries. ICP was first proposed and