Geometry of Projection

Jitendra Malik

1 Geometry of image formation

1.1 Spherical and Planar perspective

1. In spherical perspective, the projection surface is a sphere of unit radious centered at the center of projection. A point (ρ, θ, ϕ) gets mapped to $(1, \theta, \phi)$. The ray from the pint in the scene through the center of projection is perpendicular to the imaging surface.

   Spherical notation is extremely useful because it removes "position effect". Clearly same information as planar projection.

2. Vector notation A world point $\mathbf{X}$ projects to a point on a plane, $\mathbf{p}$, under planar projection, and equivalently a point on a sphere, $\mathbf{q}$ under spherical projection:

$$\begin{aligned}\mathbf{p} &= f\frac{\mathbf{X}}{\mathrm{Z}} \\ \mathbf{q} &= \frac{\mathbf{X}}{||\mathbf{X}||}\end{aligned}$$

   where $\mathbf{p} = (\mathrm{x}, \mathrm{y}, \mathrm{f})^\top$, and $\mathbf{q}$ is a unit vector. This is illustrated in figure ??. Note, precisely the same information is represented in each case - namely the ray *direction* which is the most that can be recovered from an image point. Each notation has its advantages and both will be used at various points. It is straightforward to transform between $\mathbf{p}$ and $\mathbf{q}$: $\mathbf{q} = \mathbf{p}/||\mathbf{p}||$, $\mathbf{p} = \mathrm{f}\mathbf{q}/\mathrm{q_Z}$. The connection between the vector notation and a planar image is that if $\mathbf{p} = (\mathrm{x}, \mathrm{y}, \mathrm{f})^\top$, then $\mathbf{x} = (\mathrm{x}, \mathrm{y})^\top$.

1.2 Projection of lines and planes

Vanishing points and lines (for parallel lines and planes).

1. **Lines**: A line of points in 3D can be represented as $\mathbf{X} = \mathbf{A} + \lambda\mathbf{D}$, where $\mathbf{A}$ is a fixed point, $\mathbf{D}$ a unit vector parallel to the line, and λ a measure of distance along the line. As λ increases points are increasingly further away and in the limit:

$$\lim_{\lambda\to\infty} \mathbf{p} = \mathrm{f}\frac{\mathbf{A} + \lambda\mathbf{D}}{\mathrm{A_Z} + \lambda\mathrm{D_Z}} = \mathrm{f}\frac{\mathbf{D}}{\mathrm{D_Z}}$$

   i.e. the image of the line terminates in a *vanishing point* $\mathbf{x}_\mathrm{v} = (\mathrm{fD_X}/\mathrm{D_Z}, \mathrm{D_Y}/\mathrm{D_Z})$, unless the line is parallel to the image plane ($D_Z = 0$). Note, the vanishing point is unaffected (invariant to) line position, $\mathbf{A}$, it only depends on line orientation, $\mathbf{D}$. Consequently, the family of lines parallel to $\mathbf{D}$ h ave the same vanishing point.

2. **Planes**: A plane of points in 3D can be represented as $\mathbf{X}.\mathbf{N} = \mathrm{d}$ where $\mathbf{N}$ is the unit plane normal, and d the perpendicular distance of the plane from the origin. A point $\mathbf{X}$ on the plane is imaged at $\mathbf{p} = \mathrm{f}\mathbf{X}/\mathrm{Z}$. Taking the scalar product of both sides with $\mathbf{N}$ gives $\mathbf{p}.\mathbf{N} = \mathrm{f}\mathbf{X}.\mathbf{N}/\mathrm{Z} = \mathrm{fd}/\mathrm{Z}$. In the limit of points very distant:

$$\lim_{Z\to\infty} \mathbf{p}.\mathbf{N} = 0$$

   which is the equation of a plane through the origin parallel to the world plane (i.e. which has the same normal $\mathbf{N}$). The plane $\mathbf{p}.\mathbf{N} = 0$ intersects the image plane in a *vanishing line*

$$xN_X + yN_Y + fN_Z = 0$$

   Note, the vanishing line is unaffected (invariant to) plane position, d, it only depends on plane orientation, $\mathbf{N}$. All planes with the same orientation have the same vanishing line.

   Consider a line on the plane. Where is the vanishing point of the line? As an exercise, show that the vanishing points of all lines on the plane lie on the vanishing line of the plane. Thus, two vanishing points determine the vanishing line of the plane.

Consider an observer standing on the ground looking straight ahead at a telephone pole.

The observer's image will look like that shown in Fig ? Since the ground plane has surface normal $\mathbf{N} = (0, 1, 0)$, the equation of the horizon is $y = 0$.

2 Curves and surfaces

Parametric form for curves in 2d $X(t), Y(t)$, in 3D $X(t), Y(t), Z(t)$.

Level set definition of surfaces: A surface of dimension n in R^{n+1} is a non-empty subset S of R^{n+1} of the form $S = f^{-1}(c)$ where f is a smooth function $U \subset \mathrm{R}^{n+1} \to \mathrm{R}$ with the property that $\bigtriangledown f(p) \neq \mathbf{0}$. Picking $n = 1$ above defines plane curves. Picking $n = 2$ defines 2-surfaces.

Parametrized patch definition of surfaces: A parametrized *n-surface* in R^{n+1} is the image of a smooth map $\varphi : U \to \mathrm{R}^{n+1}$, where U is a connected open set in R^n which is such that its derivative $d\varphi_p$ is non-singular (has rank n) for each $p \in U$. The image

$d\varphi_p$ is the tangent space to φ corresponding to the point $p \in U$.

The shape of the surface is captured in the *shape operator*, which measures how the surface normal $\mathbf{N}$ changes as one moves in various directions in the tangent space of the surface

Consider a point P of the surface together with all the curves at the point P that lie on the surface. The tangents to those curves at the point P all lie in a common plane, which is therefore called the tangent plane to the surface at P. The perpendicular to the tangent plane at a point P of a surface is called the normal to the surface at P. The term normal sections is used to denote curves in which the planes containing the normal intersect the surface.

To describe the curvature of surfaces in numerical terms, we consider the curvatures of the normal sections at a point P of the surface. The center of curvature of such a normal section is bound to lie on the normal to the surface passing through P, because this is the normal to all the curves that are normal sections at P. By rotating a plane passing through the normal about this normal, we get all the normal sections. In the course of the rotation, the center of curvature moves along the normal in a characteristic way.

At an elliptic point, the center of cuvrature always remains on one side of the surface. Typically, the radius of curvature will assume a maximum value r_1 for one normal section s_1 and a minimum value r_2 for another normal section s_2. We define the reciprocals of these radii of curvatures $k_1 = \frac{1}{r_1}$ and $k_2 = \frac{1}{r_2}$ to be the principal curvatures at the point P; the directions of the tangents to s_1 and s_2 are called the principal directions. It can be shown that the principal directions are mutually perpendicular.

3 How surfaces project

The basic geometry is illustrated in Figure ??. A smooth surface S is mapped by central projection to a unit sphere Σ centered at the focal point. The backprojection map F from Σ to S is defined as $F(\mathbf{p}) = \mathbf{r}(\mathbf{p}) = r(\mathbf{p})\mathbf{p}$ where $\mathbf{p}$ is a unit vector from the focal point to a point on the image sphere, and $r(\mathbf{p})$ is the distance along the visual ray from the focal point through $\mathbf{p}$ to the corresponding point $\mathbf{r} = F(\mathbf{p})$ on the surface S. We consider this map for regions of the surface where the map is not singular by excluding neighborhoods containing the occluding contour. The differential of the backprojection map F_* maps tangent vectors of Σ at $\mathbf{p}$ to tangent vectors of S at $F(\mathbf{p})$.

Define the tilt direction $\mathbf{t}$ in $T_p(\Sigma)$, the tangent plane of the viewing sphere at $\mathbf{p}$, to be a unit vector in the direction of the gradient of the distance function $r(\mathbf{p})$, and the auxiliary vector $\mathbf{b} = \mathbf{p} \times \mathbf{t}$. Then $(\mathbf{t}, \mathbf{b})$ form an orthonormal basis for the tangent plane to the image sphere Σ and together with $\mathbf{p}$ constitute an orthonormal frame field on Σ. Garding shows that $\mathbf{t}$ and $\mathbf{b}$ backproject to orthogonal vectors $F_*(\mathbf{t}) = r'\mathbf{p} + r\mathbf{t}$ and $F_*(\mathbf{b}) = r\mathbf{b}$ in the tangent space $T_{F(p)}(S)$, where r' is the directional derivative of the distance r with respect to $\mathbf{t}$. Dividing $F_*(\mathbf{t})$ and $F_*(\mathbf{b})$ by their lengths gives us an orthonormal basis $(\mathbf{T}, \mathbf{B})$ of the tangent space of the surface at $F(\mathbf{p})$. The vectors $\mathbf{T}, \mathbf{B}$ along with the unit normal to the surface $\mathbf{N} = \mathbf{T} \times \mathbf{B}$ constitute an orthonormal frame field on the surface. The slant angle σ is defined to be the angle between the surface normal $\mathbf{N}$ and the viewing direction $\mathbf{p}$, so that $\cos\sigma = \mathbf{N}.\mathbf{p}$. $F_*(\mathbf{p})$ can be expressed very conveniently in terms of the $(\mathbf{t}, \mathbf{b})$ and $(\mathbf{T}, \mathbf{B})$ bases as

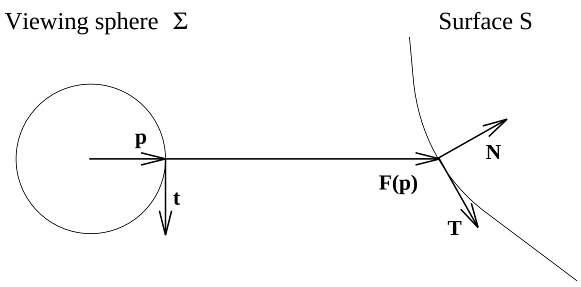

Figure 1: **Local surface geometry.**

$$F_*(\mathbf{p}) = \begin{bmatrix} r/\cos\sigma & 0 \\ 0 & r \end{bmatrix} = \begin{bmatrix} \frac{1}{m_p} & 0 \\ 0 & \frac{1}{M_p} \end{bmatrix} \quad (1)$$

where r is the distance to the object from the center of the viewing sphere, and σ is the slant angle. We see that $m_p = \cos\sigma/r$ is the scaling of the texture pattern in the "minor axis," i.e. in the tilt direction, due to projection, and $M_p = 1/r$ is the scaling in the "major axis," i.e. in the orthogonal direction. The use of the terms "major axis" and "minor axis" is intended as a mnemonic – a texture composed of circles on a planar surface would project to ellipses with minor axes aligned with the tilt direction.

4 Spatial layout, pose and shape

Spatial layout: distance and surface orientation with respect to the viewer. One way to parametrize this is using the three values distance r, slant σ, and the tilt direction. Discontinuities in distance and surface orientation are important for segmenting the scene into distinct objects.

We use two terms to describe an object

1. *Pose* how is it oriented with respect to the camera. This is measured by specifying 6 numbers (3 for its translation, 3 for rotation).

2. *Shape* : the aspect of an object's geometry that remians invariant under rotations and translation.