CSE 252A: Lenses

Lecturer: David Kriegman

Scribed by Owen Jow on October 09, 2018

1 Homographies

We can use a homography to map a point in a plane to a point in another plane under perspective projection. Let $\mathbf{x}$ be the *3D form* (in the new plane's camera frame) of a point in the original plane.

$$\mathbf{x} = \mathbf{O} + x\mathbf{i} + y\mathbf{j}$$

where x, y are the 2D coordinates of a point in the plane, $\mathbf{O}$ is the vector in 3D to the origin of the plane's coordinate frame, and $\mathbf{i}, \mathbf{j}$ are the axes in 3D of the plane.

We can also write this as a (4×3) times (3×1) matrix multiplication:

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{O} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Adding perspective projection (to bring the point into the new plane's coordinates), we have

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{f} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{O} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

This is a projective transformation (homography) from a plane to a plane. It maps planes to planes under perspective projection, taking 2D homogeneous coordinates to 2D homogeneous coordinates. Because of this, it is only defined up to a scale factor, meaning it has 8 DOF.

An `(input` x, y`)`, `(output` x', y'`)` pair consists of two measurements and gives two equations. To solve for $\mathbf{H}$, we can take 4+ inputs and 4+ outputs (yielding 8+ equations) and find a solution to $\mathbf{Ah} = \mathbf{0}$. This can be seen as solving for the eigenvector $\mathbf{h}$ corresponding to the 0 eigenvalue of $\mathbf{A}$.

We can compute a nontrivial solution for $\mathbf{H}$ using SVD. If we take the SVD of $\mathbf{A}$, $\mathbf{h}$ is given by the singular vector associated with the smallest singular value. That singular value should be close to 0.

The mapping between two images is a homography if the images have a common center of projection (which could be given by a camera rotating about one point).

2 Lenses

Real cameras aren't pinhole cameras. They have lenses, which allow us to get more light (because a lens has more space to pass through than a pinhole does) in a controlled way, without losing focus.

If we just enlarge the pinhole, it'll make everything blurry because we'll have a bunch of things mapping to each location (a point maps to a circle instead of a point). Incidentally, things also get blurry if we make the pinhole smaller due to diffraction effects around the edge of the pinhole.

Lenses increase light-gathering capacity but also try to focus light rays from a point onto a point.

2.1 Lens Models

An ideal *thin lens* is rotationally symmetric about the optical axis. Therefore, if light hits the center of projection $\mathbf{O}$, it will pass through with its direction unchanged. If light comes in parallel to the optical axis, it gets bent in such a way that all parallel lines pass through a common *focal point* $\mathbf{F}$.

Ultimately, a thin lens will make it so every ray that comes from one point will end up converging at the same point on the other side of the lens. The *thin lens equation* is

$$\frac{1}{z_{im}} + \frac{1}{z_{obj}} = \frac{1}{f}$$

for z_{im} the z-distance on the image side of the pinhole, z_{obj} the z-distance of the point in the scene, and f the focal length of the lens.

If we put the image plane at some depth, everything that converges at that depth will be imaged well. However, things which converge at different depths will be blurry. *Lenses increase light-gathering capacity, but also add blur.* In reality, only one depth is in perfect focus, although we would consider a range of depths to be in focus due to having a finite number of pixels with discrete spacing. The *depth of field* describes the range of depths that are in focus.

You can increase depth of field by shrinking the aperture. If we make the aperture smaller, the circle of confusion (the circle that points map to) at out-of-focus depths will be smaller. However, this comes at the tradeoff of less light-gathering ability.

- As photographers, we might not care about a background. We might want to make the aperture larger and just put our subject at the correct depth, s.t. the background is blurred out.

Field of view describes how much of the scene we can see in the image, and in one dimension is a function of image plane size s and focal length f – namely $\texttt{fov} = 2\tan^{-1}\left(\frac{s}{2f}\right)$. A wide-angle lens has a small f; a telephoto lens has a large f.

2.2 Deviations from Lens Models

The thin lens model is ideal and assumes the lens is infinitely thin (which it won't be in real life). In real life, we'll get distortions (aberrations). With a wide angle lens, we might get some outward-bulging (a barrel distortion), which we can correct if we know the camera parameters.

A more annoying aberration is a *spherical aberration*, where (since the lens isn't exactly rotationally symmetric) rays for a point don't exactly converge to a single point on the image side of the lens.

There's also *astigmatism*, for which there are different focal lengths for different directions over the lens and points end up converging to (oriented) lines in images, and *chromatic aberrations*, where light refracts at the lens surface according to Snell's law, which depends on the wavelength(s) of light – so different colors of light bend differently with different directions, and rays of different wavelengths converge at different depths. In an attempt to mitigate chromatic aberrations, lens makers put coatings (which are carefully engineered for refractions) on top of the glass.

Finally we have *vignetting*, where (in compound lenses) rays passing through near the outer part of the lens are blocked, and the outside of the image ends up being darker than the inside.