Perspective Projection
Problem Set 2

Computer Vision 2021
University of Bern

1 Camera models

1. **Thin lens formula**

   A camera is equipped with a lens with focal length 50mm and aperture diameter $D = 2$cm. An object at 5m from the camera is in focus. What is the minimum thickness of the camera?

2. **Thin lens formula**

   A camera is equipped with a lens with focal length 40mm and aperture diameter $D = 2$cm. The distance of the image plane to the lens plane is 5cm. What is the blur diameter of an object at infinity?

3. **Field of view**

   A simple camera is made of a CCD sensor and a single lens. The size of a CCD sensor is 4cm $\times$ 4cm. The focal length of the lens is $f = 50$mm. What is the field of view of this camera?

4. **Geometric optics**

   A simpler way to handle lenses and to determine the imaging equations is to use a *ray representation*. Instead of considering the 3D point in space, we consider a ray in space (a line) and how this is deflected by a lens when it falls within its aperture. Let us consider a lens with focal length f and a line incident at an angle θ with the lens (the angle is with respect to the normal to the lens plane) at a (signed) height d from the optical center. Compute the exiting angle ϕ by using the two basic rules of a thin lens: 1) rays through the optical center are not deflected and 2) objects at infinity are focused at a distance f behind the lens (this means that parallel rays converge at a point at a distance f behind the lens). See Fig. 1.

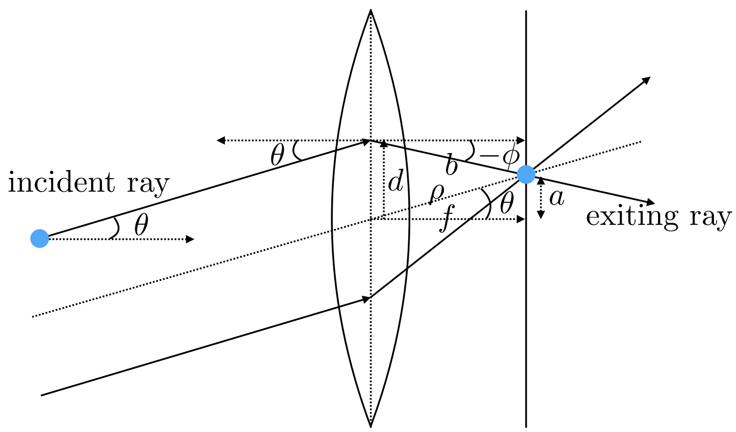

Fig. 1: Geometric optics.

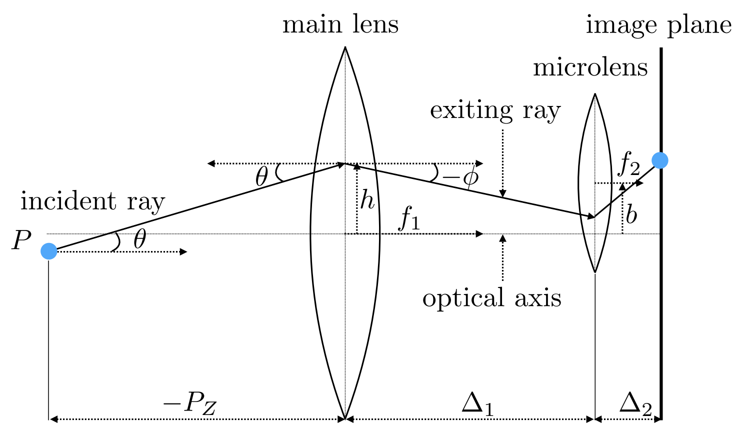

Fig. 2: Geometric optics with multiple lenses.

5. **Composite optical systems (challenging)**

   Let us consider a camera with 2 lenses and let us use the geometric optics approximation to determine the imaging equations. Suppose that the first lens is centered at the optical axis and has focal length f_1. The second lens is behind the first lens at a distance Δ_1, has focal length f_2 and its center is displaced at a height b from the optical axis. For simplicity we assume that both lenses have infinite aperture. The camera sensor is behind the second lens at a distance Δ_2. A point P is in front of the lens and emits light rays. See Fig. 2.

   Determine the ray tracing matrix M of the whole imaging system (to determine the propagation of the rays from P).

2 Filters

1. **Median filter**

   The median filter takes the middle element after sorting the elements in a patch. While this operation is quite robust to non Gaussian noise, it can lead to severe flattening of the texture when the size of the patch is large. How could we modify the median filter so that it retains the original texture while removing outliers?

2. **Nonlocal means filter (challenging)**

   The nonlocal means filter performs averaging of patches that are sufficiently similar. If we extract a patch P from the noisy image, then we select other patches P_i such that

   $$|P - P_i| < d \tag{1}$$

   with a certain scalar $d > 0$. If N patches P_i satisfy the constraint above, then the updated/denoised patch $\hat{P}$ is obtained via averaging

   $$\hat{P} = \frac{1}{N}\sum_{i=1}^{N} P_i. \tag{2}$$

   Suppose that the original image is affected by zero-mean i.i.d. Gaussian noise. Then, we can write

   $$P = P_0 + n \qquad\qquad P_i = P_0 + n_i \tag{3}$$

   where P_0 is the noise-free patch and $n, n_i \sim \mathcal{N}(0, \sigma^2 I_d)$ are Gaussian random variables. What is the distribution of $\hat{P}$?

3. **Mean shift filter (challenging)**

   The *mean shift* filter is a procedure to find the maxima of a probability density function (the modes) from its samples. The first step is to assume that the probability density function can be approximated by its kernel density estimate

   $$p(x) = \frac{1}{N}\sum_{i=1}^{N} K\left(\frac{|x - x_i|^2}{h}\right) \tag{4}$$

   where $K(|x|^2)$ is a kernel that integrates to 1 in x and h is a bandwidth parameter. A necessary condition to determine the maxima of a function is that the first order derivatives must be zero. By taking the first order derivatives in x we obtain

   $$\nabla p(x) = \frac{1}{N}\sum_{i=1}^{N} 2\frac{x - x_i}{h} K'\left(\frac{|x - x_i|^2}{h}\right) = 0 \tag{5}$$

Next we rearrange the equation above to obtain the following iterative update

$$x^{t+1} = \frac{\sum_{i=1}^{N} x_i K'\left(\frac{|x^t - x_i|^2}{h}\right)}{\sum_{i=1}^{N} K'\left(\frac{|x^t - x_i|^2}{h}\right)} \tag{6}$$

This update is the *mean shift* algorithm and it can be shown to converge to the modes of the distribution $p(x)$ in a number of cases. Let us consider the Epanechnikov kernel

$$K(z) = \begin{cases} \frac{3}{4}(1-z) & 0 \leq z \leq 1 \\ 0 & z > 1 \end{cases} \tag{7}$$

then, we have

$$K'(z) = \begin{cases} -\frac{3}{4} & 0 \leq z \leq 1 \\ 0 & z > 1 \end{cases} \tag{8}$$

and the meanshift iteration becomes

$$x^{t+1} = \frac{\sum_{i=1}^{N} x_i \delta(|x^t - x_i|^2 < h)}{\sum_{i=1}^{N} \delta(|x^t - x_i|^2 < h)} \tag{9}$$

where $\delta(\text{event})$ is the indicator function of “event” (it is 1 if the event is true and 0 otherwise). Find the relation between the mean shift filter with Epanechnikov kernel and the nonlocal mean filter above.

3 Edges

1. **Gradients**

   Let us consider an image x. Let also $p, q \in \mathbf{R}^2$ be two pixels (represented as two 2D vectors) in x. We define *self-similarity* as the property that

   $$x[q] = x[\alpha(q-p)+p] \qquad \forall q : |q-p| \leq \rho \tag{10}$$

   where $\rho > 0$ is the radius of a ball around p and $0 < \alpha < 1$. See Fig. 3. In other words, rays originating at p should have constant image intensity. Notice that, in particular, the self-similarity property is satisfied at corners and at edges. How does the gradient $\nabla x[q]$ in the point q relate to the self-similarity in the limit as $\alpha \to 1$?

2. **Gradients**

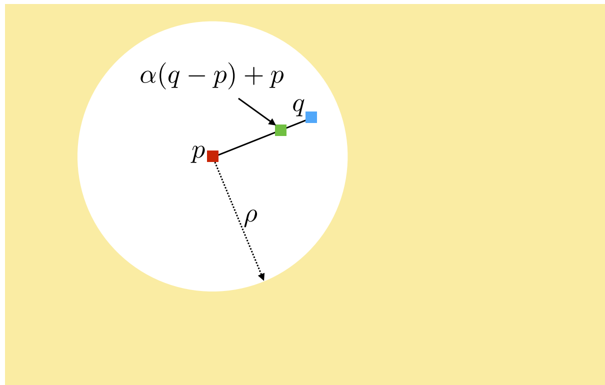

Fig. 3: Gradients and self-similarity.

Show that the family of all images x that can be written as a function of only the direction of $q-p$

$$x[q] = f\left(\frac{q-p}{|q-p|}\right) \tag{11}$$

for any function f, satisfies the property below for a fixed pixel p

$$\nabla x[q]^T(q-p) = 0 \qquad \forall q. \tag{12}$$