# Multiple View Geometry: Solution Sheet 4

Prof. Dr. Daniel Cremers, Nikolaus Demmel, Marvin Eisenberger
Computer Vision Group, TU Munich
`http://vision.in.tum.de/teaching/ss2018/mvg2018`

Exercise: May 16th, 2018

## Part I: Theory

1. **Image Formation**

   (a) Compute $\lambda$ and show that (2) is equivalent to

   $$u = \frac{fX}{Z} + o_x , \quad v = \frac{fY}{Z} + o_y .$$

   Performing the matrix multiplication in (2), one obtains

   $$\begin{pmatrix} \lambda u \\ \lambda v \\ \lambda \end{pmatrix} = \begin{pmatrix} fX + o_x Z \\ fY + o_y Z \\ Z \end{pmatrix}$$

   From the third row, it directly follows that $\lambda = Z$. Substituting $Z$ for $\lambda$ and dividing the equation by $Z$, one immediately obtains the result.

   (b) A classic ambiguity of the perspective projection is that one cannot tell an object from another object that is exactly *twice as big but twice as far*. Explain why this is true.
   Let $\tilde{\mathbf{X}}_1 = (X_1\ Y_1\ Z_1)^\top$ be a point on the smaller object and $\tilde{\mathbf{X}}_2 = (X_2\ Y_2\ Z_2)^\top$ a point on the larger object. Since $\tilde{\mathbf{X}}_2$ is twice as far away, we have $Z_2 = 2Z_1$, and since it is twice as big we have $X_2 = 2X_1$ and $Y_2 = 2Y_1$. Thus,

   $$u_2 = \frac{fX_2}{Z_2} + o_x = \frac{2fX_1}{2Z_1} + o_x = \frac{fX_1}{Z_1} + o_x = u_1$$

   and analogous for $v_2 = v_1$.

   (c) For a camera with $f = 540$, $o_x = 320$ and $o_y = 240$, compute the pixel coordinates $u$ and $v$ of a point $\tilde{\mathbf{X}} = (60\ 100\ 180)^\top$.

   $$u = \frac{fX}{Z} + o_x = \frac{540 \cdot 60}{180} + 320 = 500$$

   $$v = \frac{fY}{Z} + o_y = \frac{540 \cdot 100}{180} + 240 = 540$$

   Explain with the help of (b) why the units of $\tilde{\mathbf{X}}$ are not needed for this task.
   Using different units (mm, cm, m, etc.) can be interpreted as scaling the point coordinates by a constant factor (10, 100, ...). The argument of (b) for a factor of 2 can easily be generalized to any factor $\alpha$.
   Will the projected point be in the image if it has dimensions $640 \times 480$?
   No, the point $(u, v) = (500, 540)$ is not in $[-0.5, 639.5] \times [-0.5, 479.5]$.

(d) Using the generic projection $\pi$, show that (3) — and therefore also (1) and (2) — is equivalent to

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K \begin{pmatrix} \pi(\tilde{\mathbf{X}}) \\ 1 \end{pmatrix}.$$

Insert in the RHS of the equation:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K \begin{pmatrix} \pi(\tilde{\mathbf{X}}) \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & o_x \\ 0 & f & o_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X/Z \\ Y/Z \\ 1 \end{pmatrix} = \begin{pmatrix} fX/Z + o_x \\ fY/Z + o_y \\ 1 \end{pmatrix}$$

2. **Radial Distortion**

(a) Can this model be used for lenses with a field of view of more than $180°$?

No, it can only model points for which the viewing ray intersects the image plane. Since points are first projected on the canonical iamge plane with $\pi(\tilde{\mathbf{X}})$ there is a singularity at $Z = 0$ (which is $180°$).

*Note:* It is possible to rewrite the FOV model to avoid the division by $Z$ and apply it to lenses with more than $180°$ FOV.

(b) Derive a closed form solution for $f$ in the undistortion formula

$$\pi(\tilde{\mathbf{X}}) = f\left(\|\pi_d(\tilde{\mathbf{X}})\|\right) \cdot \pi_d(\tilde{\mathbf{X}})$$

using (4) and $g(r) = g_{\text{ATAN}}(r)$.

Define $r := \|\pi(\tilde{\mathbf{X}})\|$ and $r_d := \|\pi_d(\tilde{\mathbf{X}})\|$. The norms of (4) and (6) are:

$$r = f(r_d)r_d \quad \text{and} \quad r_d = g(r)r$$

Inserting $g = g_{\text{ATAN}}$ yields

$$r_d = \frac{1}{\omega r} \arctan\left(2r \tan\left(\frac{\omega}{2}\right)\right) \quad r = \frac{1}{\omega} \arctan\left(2r \tan\left(\frac{\omega}{2}\right)\right)$$

$$\Rightarrow \tan(r_d \omega) = 2r \tan\left(\frac{\omega}{2}\right)$$

$$\Rightarrow r = \frac{\tan(r_d \omega)}{2 \tan\left(\frac{\omega}{2}\right)} = f(r_d)r_d \quad \Rightarrow f(r_d) = \frac{\tan(r_d \omega)}{2 r_d \tan\left(\frac{\omega}{2}\right)}$$

3. **Image Rectification**

Let $I_d$ be a distorted image taken by a camera with intrinsic parameter matrix $K_d$.

Write down expressions for the pixel coordinates $u_d$, $v_d$ of a point $\tilde{\mathbf{X}} \in \mathbb{R}^3$ in the distorted image and for the pixel coordinates $u$, $v$ of the same point projected by a perfect pinhole camera with intrinsic camera matrix $K$.

$$\begin{pmatrix} u_d \\ v_d \\ 1 \end{pmatrix} = K_d \begin{pmatrix} \pi_d(\tilde{\mathbf{X}}) \\ 1 \end{pmatrix} = K_d \begin{pmatrix} g\left(\|\pi(\tilde{\mathbf{X}})\|\right)\pi(\tilde{\mathbf{X}}) \\ 1 \end{pmatrix}, \quad \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K \begin{pmatrix} \pi(\tilde{\mathbf{X}}) \\ 1 \end{pmatrix}$$