Unified Omnidirectional Camera Model

Luca Di Giammarino

Draft

The pinhole model is simple and accurate, above all is a linear model. That means that world lines are perceived as lines in the image too. However, lenses always produce some distortion and lines in the world are always perceived as curves in the image. We have two possible solutions for distortion. One is to rectify the image and the other is to use a more suitable camera model. In order to rectify the image, one need to estimate some sort of transformation that allows to correct the image measurements to those that would have been obtained under a perfect linear camera action (most famous distortion model are radial and tangential, see Hartley and Zisserman for more details). The drawbacks of rectifying an image is that we loose most of our field of view as we can see in Fig. 1.

We can avoid rectification by using a more suitable camera model (i.e. unified omnidirectional camera model). The major advantages of this is that it can accurately model the geometric image formation for a wide range of imaging devices and lenses. Differently from pinhole model, a omnidirectional camera model has five parameters (one more compared to pinhole), focal length f_x, f_y, camera centers c_x, c_y and the distance between camera center and unit sphere center ξ. A 3D point in Euclidean camera coordinates is first projected onto a camera centered unit sphere (see Fig. 2). Then the point is projected to an image plane as in the pinhole model through a center with an offset $-\xi$ along the z axis. The projection of point is computed as follows:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f_x \frac{x}{z+||\mathbf{p}_w||\xi} \\ f_y \frac{y}{z+||\mathbf{p}_w||\xi} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix}$$

With $\mathbf{p}_w = \begin{bmatrix} x & y & z \end{bmatrix}^T$ denoting our 3D point in world. Note that, if $\xi = 0$ the unit sphere center coincide with the origin of the camera reference

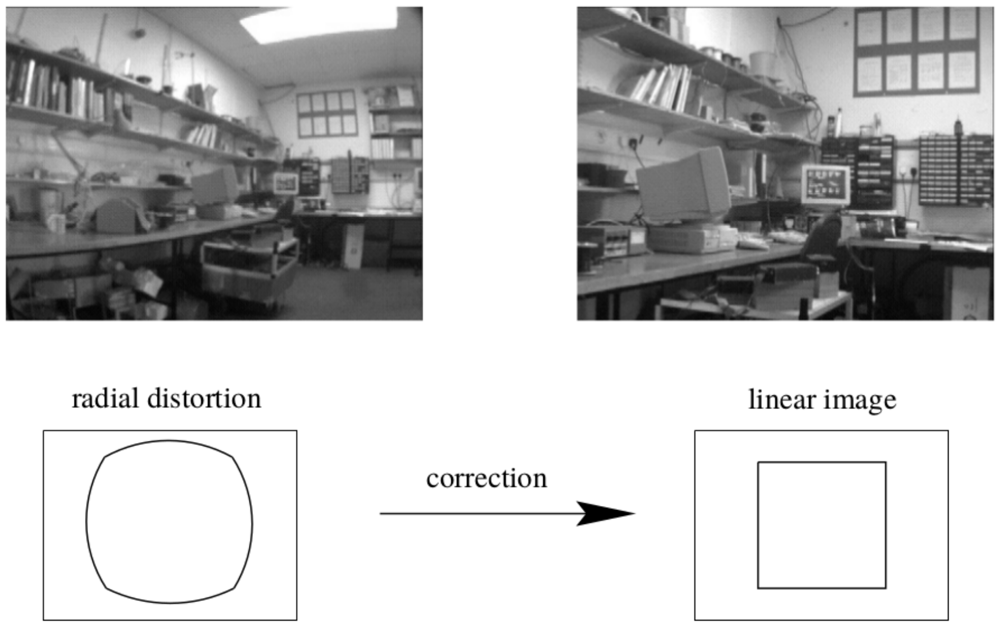

Figure 1: **Distorted and undistorted images**. Distorted (left) and undistorted (right) images, note the curved image lines in (left) which in reality are straight world lines. Down we can see geometrically the difference between a square with radial distortion, corrected to one that would have been obtained under a perfect linear lens. Courtesy of Hartley and Zisserman.

frame, we degenerate to classic pinhole model. We can write the unified projection embedding the transformation matrix $\mathbf{X}$, that maps worlds points onto the camera reference frame $\mathbf{p}_c = \mathbf{X}\mathbf{p}_w$. While we denote with $\mathbf{p}_s = \begin{bmatrix} x_s & y_s & z_s \end{bmatrix}^T = \frac{\mathbf{p}_c}{\|\mathbf{p}_c\|}$ the point in unit sphere (direction). Hence we have:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \mathbf{K}\mathrm{proj}_u\left(\frac{\mathbf{p}_c}{\|\mathbf{p}_c\|}\right)$$

$$\mathrm{proj}_u(\mathbf{p}_s) = \begin{bmatrix} \frac{x_s}{z_s+\xi} & \frac{y_s}{z_s+\xi} & 1 \end{bmatrix}^T$$

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \end{bmatrix}$$

Note that in $\mathbf{K}$ we drop last row (compared to classic pinhole intrinsics matrix) to be compliant with matrix multiplication, since for the unified model $\mathbf{K}$ appears after homogenous division (in pinhole model is the opposite).

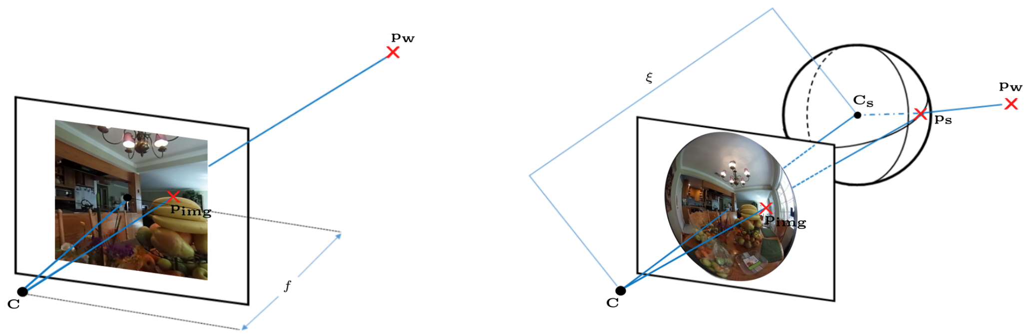

Figure 2: **Pinhole model and omnidirection unified model**. On the left the standard pinhole projection and on the right the unified omnidirectional camera model projection. In this last the image coordinate of a 3D point $\mathbf{p}_w$ is found by first projecting it on the unit sphere, and then projecting it to image plane. The distance between the camera center and the unit sphere center is represented by ξ, the focal length f parameterize always the offset between the camera center and the image plane (not graphically represented on the right image for simplicity).

We can finally compute the Jacobian, decompose it using the chain rule in $\mathbf{J}_{\text{icp}}$, $\mathbf{J}_{\text{normalize}}$ and $\mathbf{J}_{\text{proj}_u}$ as follows

$$\frac{\partial \mathbf{e}^{[n,m]}(\mathbf{X} \boxplus \mathbf{\Delta x})}{\partial \mathbf{\Delta x}}\Bigg|_{\mathbf{\Delta x}=\mathbf{0}} = \mathbf{K} \underbrace{\frac{\partial \text{proj}_u(\mathbf{p})}{\partial \mathbf{p}}}_{\mathbf{J}_{\text{proj}_u}}\Bigg|_{\mathbf{p}=\hat{\mathbf{p}_s}^{[n]}} \underbrace{\frac{\partial \frac{\mathbf{p}}{\|\mathbf{p}\|}}{\partial \mathbf{p}}}_{\mathbf{J}_{\text{normalize}}}\Bigg|_{\mathbf{p}=\hat{\mathbf{p}_c}^{[n]}} \underbrace{\frac{\partial \mathbf{h}_{\text{icp}}^{[n]}(\mathbf{X} \boxplus \mathbf{\Delta x})}{\partial \mathbf{\Delta x}}}_{\mathbf{J}_{\text{icp}}}\Bigg|_{\mathbf{\Delta x}=\mathbf{0}}$$

$$= \mathbf{K}\mathbf{J}_{\text{proj}_u}(\hat{\mathbf{p}}_s^{[n]})\mathbf{J}_{\text{normalize}}(\hat{\mathbf{p}}_c^{[n]})\mathbf{J}_{\text{icp}}^{[n]}(\mathbf{p}_w)$$

$$\mathbf{J}^{[n]}_{\text{icp}} = \begin{bmatrix} \mathbf{I}_{3\times3} & \lfloor \mathbf{R}\mathbf{p}^{(n)}_w + \mathbf{t} \rfloor_\times \end{bmatrix}$$

$$\mathbf{J}^{[n]}_{\text{normalize}} = \frac{1}{\|\mathbf{p}_c\|^3} \left[\mathbf{I}_{3\times3} \|\mathbf{p}_c\|^2 - \mathbf{p}_c \mathbf{p}_c^T \right]$$

$$\mathbf{J}^{[n]}_{\text{proj}_u} = \begin{bmatrix} \frac{1}{z_s+\xi} & 0 & -\frac{x_s}{(z_s+\xi)^2} \\ 0 & \frac{1}{z_s+\xi} & -\frac{y_s}{(z_s+\xi)^2} \end{bmatrix}$$

The drawbacks of this unified model is that you need to develop the epipolar geometry yourself, since important computer vision books treat only pinhole model.