Arriving at $Z/f$

Michael Hornáček
CVPR 2014 Submission #1678
Supplementary Material

We state in Section 2.1 of the paper that it is easy to show that the distance between two points both situated at the same depth $Z$ and projecting to neighboring pixels is given by $Z/f$, where $f$ is the camera’s focal length in units of pixels (cf. Figure 1). We motivate this result by computing the distance between the point $P$ at depth $Z$ projecting to the pixel $(i,j)^\top$, and the points at the same depth $Z$ projecting to the pixel neighbors $(i+1,j)^\top, (i,j+1)^\top$, respectively. The point $P$ is given by $Z \cdot K^{-1} (i,j,1)^\top$, where $K$ is the $3 \times 3$ camera calibration matrix (cf. Hartley and Zisserman [1]):

$$
\begin{bmatrix}
  f & 0 & px \\
  0 & f & py \\
  0 & 0 & 1
\end{bmatrix},
$$

where $(p_x, p_y)^\top$ is the camera’s principal point in pixel units. It follows that $P$ is given in closed form by

$$
Z \cdot K^{-1} \begin{bmatrix} i \\ j \\ 1 \end{bmatrix} = Z \begin{bmatrix} 1/f & 0 & -p_x/f \\ 0 & 1/f & -p_y/f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} i \\ j \\ 1 \end{bmatrix} = Z \begin{bmatrix} \frac{i-p_x}{f} \\ \frac{j-p_y}{f} \\ 1 \end{bmatrix}.
$$

**Horizontal Neighbors.** We first consider the Euclidean distance $\delta_Z$ between two points both situated at depth $Z$ and projecting to the horizontal pixel neighbors $(i,j)^\top, (i+1,j)^\top$:

$$
\delta_Z = \| Z \cdot K^{-1} (i+1,j,1)^\top - Z \cdot K^{-1} (i,j,1)^\top \|_2
= \| Z \left( K^{-1} (i+1,j,1)^\top - K^{-1} (i,j,1)^\top \right) \|_2
= \left\| Z \begin{bmatrix} \frac{i+1-p_x}{f} - \frac{i-p_x}{f} \\ \frac{j-p_y}{f} - \frac{j-p_y}{f} \\ 1 - 1 \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} Z \frac{i}{f} \\ 0 \\ 0 \end{bmatrix} \right\|_2 = \sqrt{\left( Z \frac{i}{f} \right)^2} = Z/f.
$$
Figure 1: The Euclidean distance $Z/f$ in world space between two points $P_x, P_{x'}$ both situated at depth $Z$—with respect to the camera center $C$—and projecting to neighboring pixels $x, x'$.

**Vertical Neighbors.** Finally, we consider the analogous Euclidean distance $\delta_Z'$ for the vertical pixel neighbors $(i, j)^T, (i, j + 1)^T$:

$$
\delta_Z' = \| Z \cdot K^{-1}(i, j + 1, 1)^T - Z \cdot K^{-1}(i, j, 1)^T \|_2
= \| Z (K^{-1}(i, j + 1, 1)^T - K^{-1}(i, j, 1)^T) \|_2
= \| Z \left( \frac{i-p_x}{f} - \frac{i-p_x}{j+1-p_y} \right) \|_2
= \| \begin{pmatrix} 0 \\ \frac{Z}{f} \end{pmatrix} \|_2
= \sqrt{\left( \frac{Z}{f} \right)^2} = Z/f.
$$

**References**