Supplementary material for Omobot: a low-cost mobile robot for autonomous search and fall detection

Shihab Uddin Ahamad*, Masoud Ataei*, Vijay Devabhaktuni†, Vikas Dhiman*
*Electrical and Computer Engineering Department,
University of Maine, Orono, ME, USA
shihab.ahamad@maine.edu
†Electrical Engineering Department,
Illinois State University, Normal, IL, USA

Appendix A
Homography from equation of a plane

Proof. From pinhole camera projection equations for camera 1 and 2 we have [1],

$$\lambda_1 \mathbf{u}_1 = K_1 \mathbf{x}_1, \qquad \lambda_2 \mathbf{u}_2 = K_2 \underbrace{({}^2R_1 \mathbf{x}_1 + {}^2\mathbf{t}_1)}_{\mathbf{x}_2}, \tag{1}$$

where $\lambda_1, \lambda_2 \in \mathbb{R}$ are scalars that represent the depth of $\mathbf{x}_1$ from the respective camera origins and $\mathbf{x}_2 \in \mathbb{R}^3$ are the coordinates of the point $\mathbf{x}_1$ in the frame of camera 2. We can solve for λ_1 by computing the intersection of the ray $\mathbf{x}_1 = \lambda_1 K_1^{-1}\mathbf{u}_1$, with the plane $\mathcal{P}$.

$$\hat{\mathbf{n}}^\top(\lambda_1 K_1^{-1}\mathbf{u}_1) = h \implies \lambda_1 = \frac{h}{\hat{\mathbf{n}}^\top K_1^{-1}\mathbf{u}_1} \tag{2}$$

Now, substituting λ_1 in (1) and rearrange the equation to form the H matrix:

$$\begin{aligned} \lambda_2 \mathbf{u}_2 &= \frac{h}{\hat{\mathbf{n}}^\top K_1^{-1}\mathbf{u}_1} K_2 {}^2R_1 K_1^{-1}\mathbf{u}_1 + K_2 {}^2\mathbf{t}_1 \\ (\hat{\mathbf{n}}^\top K_1^{-1}\mathbf{u}_1)\lambda_2\mathbf{u}_2 &= hK_2 {}^2R_1 K_1^{-1}\mathbf{u}_1 + K_2 {}^2\mathbf{t}_1(\hat{\mathbf{n}}^\top K_1^{-1}\mathbf{u}_1) \\ \underbrace{(\hat{\mathbf{n}}^\top K_1^{-1}\mathbf{u}_1\lambda_2)}_{\alpha}\mathbf{u}_2 &= \underbrace{(hK_2 {}^2R_1 K_1^{-1} + K_2 {}^2\mathbf{t}_1\hat{\mathbf{n}}^\top K_1^{-1})}_{{}^2H_1}\mathbf{u}_1. \end{aligned} \tag{3}$$

In the last step, we can absorb $\hat{\mathbf{n}}^\top K_1^{-1}\mathbf{u}_1\lambda_2$ in α as a scaling factor because α can absorb an arbitrary non-zero scalar function of $\mathbf{u}_1$.

To justify this possibility of α, let us rewrite $\alpha\mathbf{u}_2 = {}^2H_1\mathbf{u}_1$ with α as a multiple of arbitrary non-zero scalar function $f(\mathbf{u}_1) : \mathbb{P}^2 \to (\mathbb{R} \setminus \{0\})$ of $\mathbf{u}_1$ with $\alpha = \gamma f(\mathbf{u}_1)$ where $\gamma \in \mathbb{R}$ is a scale factor,

$$\gamma f(\mathbf{u}_1)\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{h}_1^\top \\ \mathbf{h}_2^\top \\ \mathbf{h}_3^\top \end{bmatrix}\mathbf{u}_1 = {}^2H_1\mathbf{u}_1. \tag{4}$$

Solving for γ we can equate, $\gamma f(\mathbf{u}_1) = \mathbf{h}_3^\top\mathbf{u}_1$ which gives, $\gamma = \frac{\mathbf{h}_3^\top\mathbf{u}_1}{f(\mathbf{u}_1)}$. Substitute γ in (4), to get the original definition of homography transform,

$$\frac{\mathbf{h}_3^\top\mathbf{u}_1}{f(\mathbf{u}_1)} f(\mathbf{u}_1)\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \underbrace{\mathbf{h}_3^\top\mathbf{u}_1}_{\alpha'}\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = {}^2H_1\mathbf{u}_1. \tag{5}$$

□

When we implement the homography on the robot and in practice use (3), we made these assumptions:

- All the points of detected human lies on a plane
- The plane is parallel to x-y axis of the camera, then $\hat{\mathbf{n}}^\top = [0., 0., 1.]$
- The person distance is in the range of 1 to 3 meters and we transform the image based on the different expected distance range from LiDAR estimation, for instance, the H matrix for $h = 3$ is:

$${}^2H_1 = \begin{bmatrix} 0.896 & -0.219 & 13.76 \\ 0.000 & 0.683 & 16.55 \\ 0.000 & -0.002 & 1.000 \end{bmatrix} \tag{6}$$

References

[1] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. New York, NY, USA: Cambridge University Press, 2003.