Supplementary document

1 Camera calibration

We sought to convert image frame coordinates in pixels to real-world measurements in cm by calibrating the camera so that its position and orientation with respect to room exit is known accurately.

We denote the camera frame by $\mathcal{C}$ and the world frame by $\mathcal{W}$. The world frame in the camera frame is denoted by the transformation ${}^{\mathcal{C}}T_{\mathcal{W}} \in SE(3)$. A point in the world frame ${}^{\mathcal{W}}\boldsymbol{r} \in \mathbb{R}^3$ is converted to camera frame by the following transformation [3].

$$ {}^{\mathcal{C}}\tilde{\boldsymbol{r}} = {}^{\mathcal{C}}T_{\mathcal{W}} {}^{\mathcal{W}}\tilde{\boldsymbol{r}} \tag{1} $$

where ${}^{\mathcal{W}}\tilde{\boldsymbol{r}} = \begin{bmatrix} {}^{\mathcal{W}}\boldsymbol{r}^{\mathrm{T}} & 1 \end{bmatrix}^{\mathrm{T}}$ denotes the homogeneous coordinate. The same location on the camera frame is converted to pixel values (u, v) on the camera image by multiplying it to the camera calibration matrix

$$ K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} $$

so that

$$ \begin{aligned} \begin{bmatrix} u {}^{\mathcal{C}}r_3 \\ v {}^{\mathcal{C}}r_3 \\ {}^{\mathcal{C}}r_3 \end{bmatrix} &= K \begin{bmatrix} I | \mathbf{0}_{3\times 1} \end{bmatrix} {}^{\mathcal{C}}\tilde{\boldsymbol{r}} \\ &= K \begin{bmatrix} I | \mathbf{0}_{3\times 1} \end{bmatrix} {}^{\mathcal{C}}T_{\mathcal{W}} {}^{\mathcal{W}}\tilde{\boldsymbol{r}} \\ &= \boldsymbol{P} {}^{\mathcal{W}}\tilde{\boldsymbol{r}} \end{aligned} \tag{2} $$

where $\boldsymbol{P} \in \mathbb{R}^{3\times 4}$ is the projection matrix. The pixel location can be calculated by dividing the left hand side by ${}^{\mathcal{C}}r_3$.

In order to convert the camera pixel coordinates into a world frame with the origin near the exit door, the following steps were performed:

1. We first calibrated the camera for intrinsic parameters (focal length and location of centroid along each axis) using the MATLAB calibration method [1]. This gave us the camera calibration matrix K.
2. We then manually marked points on the image of the floor of the experimental room whose exact location was known with respect to the exit location. Specifically, we marked the corners of the square tiles on the floor which were 54.5 cm wide. We marked 20 such points on a 4 × 5 grid.

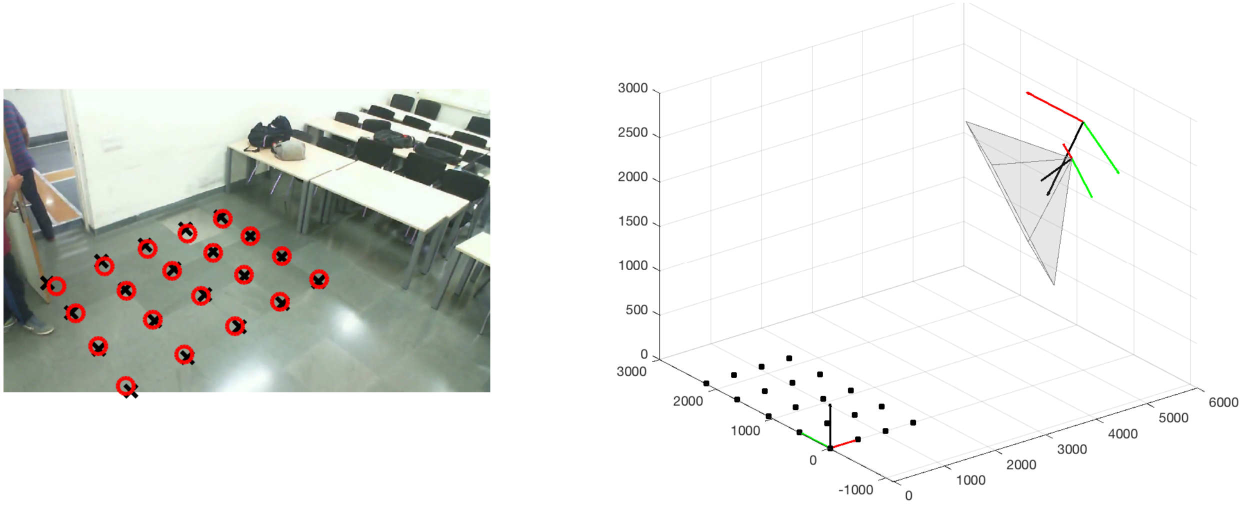

Figure 1: Reprojected points on the floor (left) after estimating the camera extrinsic parameters (right) for a select trial. The coordinate frames show the initial estimate (no gray pyramid showing field-of-view) and the final estimate (gray pyramid showing camera field of view)

3. We then initialized the camera location at approximately 2.7 m high above the floor and oriented towards the center of the room, and minimized the sum of distance between the reprojected three-dimensional locations on the floor and the marked locations on the image over a range of camera extrinsic parameters (position and orientation). The values that minimized this quantity was used to record the camera extrinsic calibration.

4. With the assumption that the average height of all participants (${}^{\mathcal{W}}r_3$) was 1.65 m [2], we then computed the three-dimensional location of each tracked point in the video to obtain the world coordinates. Specifically, we assumed that. Plugging this into equation (2), we solve for the three unknowns: ${}^{\mathcal{W}}r_1, {}^{\mathcal{W}}r_2, {}^{\mathcal{C}}r_3$

$$\begin{aligned} \begin{bmatrix} u\,{}^{\mathcal{C}}r_3 \\ v\,{}^{\mathcal{C}}r_3 \\ {}^{\mathcal{C}}r_3 \end{bmatrix} &= \boldsymbol{P} \begin{bmatrix} {}^{\mathcal{W}}r_1 \\ {}^{\mathcal{W}}r_2 \\ {}^{\mathcal{W}}r_3 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} P_{11}{}^{\mathcal{W}}r_1 + P_{12}{}^{\mathcal{W}}r_2 + P_{13}{}^{\mathcal{W}}r_3 + P_{14} \\ P_{21}{}^{\mathcal{W}}r_1 + P_{22}{}^{\mathcal{W}}r_2 + P_{23}{}^{\mathcal{W}}r_3 + P_{24} \\ P_{31}{}^{\mathcal{W}}r_1 + P_{32}{}^{\mathcal{W}}r_2 + P_{33}{}^{\mathcal{W}}r_3 + P_{34} \end{bmatrix} \end{aligned} \tag{3}$$

which can be rearranged so that

$$\begin{bmatrix} {}^{\mathcal{C}}r_3 \\ {}^{\mathcal{W}}r_1 \\ {}^{\mathcal{W}}r_2 \end{bmatrix} = \begin{bmatrix} u & -P_{11} & -P_{12} \\ v & -P_{21} & -P_{22} \\ 1 & -P_{31} & -P_{32} \end{bmatrix}^{-1} \begin{bmatrix} P_{13}{}^{\mathcal{W}}r_3 + P_{14} \\ P_{23}{}^{\mathcal{W}}r_3 + P_{24} \\ P_{33}{}^{\mathcal{W}}r_3 + P_{34} \end{bmatrix} \tag{4}$$

2 Experimental data

Table 1: Experiments performed with group ids. Every group performed at most two trials, and no group was repeated for the same condition.

rush (%)	**no-rush (%)**	**# trials**	**Group id (size)**
100	0	2	C (24), F (21)
75	25	2	A (21), B (22)
50	50	1	A (21)
25	75	2	D (23), E (22)
0	100	2	B (22), C (24)

Data for the seven trials that were considered for this paper can be downloaded from: `https://www.dropbox.com/s/kwb1rl3b616w2od/supplementary_data_for_paper.xlsx?dl=0`

	Name	**Estimate**	**SE**	**DF**	p	**Lower Bound**	**Upper Bound**
exit speed (Rush)	Intercept	0.99597	0.085155	99	< 0.01	0.827	1.1649
	fraction of rush	-0.34038	0.10426	99	< 0.01	-0.54726	-0.1335
exit speed (No-Rush)	Intercept	0.57151	0.059059	96	< 0.01	0.45428	0.68875
	fraction of rush	-0.0078569	0.14656	96	0.95736	-0.29877	0.28305
	Name	**Estimate**	**SE**	**DF**	p	**Lower Bound**	**Upper Bound**
deviation rate (Rush)	Intercept	9.2817	2.538	99	< 0.01	4.2457	14.318
	fraction of rush	6.3946	3.4303	99	0.06	-0.41174	13.201
deviation rate (No-Rush)	Intercept	13.774	1.8247	96	< 0.01	10.152	17.396
	fraction of rush	4.9671	4.3258	96	0.25371	-3.6194	13.554

Table 2: Linear mixed-effects models for exit speed and deviation rate after including the two rejected trials

References

[1] Jean-Yves Bouguet. Camera Calibration Toolbox for Matlab, 2010.

[2] A. Deaton. Height, health, and inequality: the distribution of adult heights in india. *American Economic Review*, 98(2):468–74, 2008.

[3] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision.* Cambridge university press, 2003.