Plane and Edge Point based SLAM

Sun Qinxuan

July 19, 2017

1 Introduction

2 System Overview

see Section 4.1.1 for details;

3 Constraint Analysis

4 Plane and Point based Camera Tracking with Known Correspondences

There are N_π pairs of matched planes $\{{}^c\pi_i, {}^r\pi_i\}_{i=1,\cdots,N}$ between two frames (or between the map and one frame). $\left\{{}^c\mathbf{p}^i_{s,k}, {}^c\mathbf{p}^i_{o,k}\right\}_{k=1,\cdots,N^i_p}$ denote shadow points ${}^c\mathbf{p}^i_{s,k}$ on plane ${}^c\pi_i$ with corresponding occluding points ${}^c\mathbf{p}^i_{o,k}$ causing the shadow.

$\mathbf{w} = [\mathbf{t}^T, \omega^T]^T$ denotes the 6-DoF camera pose.

i	plane index;
k	point index;
c presuperscript	current frame;
r presuperscript	reference frame;
$\pi_i = [\mathbf{n}_i^T, d_i]^T$	parameters of the i-th plane;
$\mathbf{p}_{o,k}$	the k-th occluding edge point;
$\{{}^c\pi_i, {}^r\pi_i\}_{i=1,\cdots,N}$	matched planes;
$\{{}^c\mathbf{p}_{o,k}, {}^r\mathbf{p}_{o,k}\}_{k=1,\cdots,N_p}$	matched occluding edge points;
$\mathbf{w} = [\mathbf{t}^T, \omega^T]^T \in \mathbb{R}^6$	6-DoF camera pose;
$T_{cr} \in \mathbb{SE}(3)$	transformation from reference to current frame;
$\mathbf{R}_{cr} \in \mathbb{SO}(3)$	rotation matrix;
$\mathbf{t}_{cr} \in \mathbb{R}^3$	translation vector;

4.1 Plane based Tracking

4.1.1 Plane Parameters' Constraints on Camera Motion

Suppose that the camera motion $\mathbf{w}$ is calculate such that the corresponding planes in two successive frames coincide with each other. The objective function is defined as

$$J_\pi(\mathbf{w}) = \sum_{i=1}^{N} J_{\pi,i} = \sum_{i=1}^{N} \left(\|{}^c\mathbf{n}_i - \mathbf{R}_{cr}{}^r\mathbf{n}_i\|^2 + \left[{}^c d_i - \left({}^r d_i + {}^c\mathbf{n}_i^T \mathbf{t}_{cr}\right)\right]^2 \right) \tag{1}$$

The matrix $\mathbf{\Psi}_\pi$ is calculated by

$$\begin{aligned} \mathbf{\Psi}_\pi &= \sum_{i=1}^{N} \left(\frac{\partial J_{\pi,i}}{\partial \mathbf{w}}\right) \left(\frac{\partial J_{\pi,i}}{\partial \mathbf{w}}\right)^T \\ &= \sum_{i=1}^{N} 4 \begin{bmatrix} \left({}^r d_i - {}^c d_i\right)^2 {}^c\mathbf{n}_i {}^c\mathbf{n}_i^T & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \left({}^c\mathbf{n}_i \times {}^r\mathbf{n}_i\right)\left({}^c\mathbf{n}_i \times {}^r\mathbf{n}_i\right)^T \end{bmatrix} \end{aligned} \tag{2}$$

The matrix $\mathbf{\Psi}_\pi$ is actually a scatter matrix which contains information about the distribution of the gradient of $J_{\pi,i}$ w.r.t. $\mathbf{w}$ over all planes in the matched plane set. Performing principal component analysis upon $\mathbf{\Psi}_\pi$ results in

$$\mathbf{\Psi}_\pi = Q_\pi \Lambda_\pi Q_\pi^T = \begin{bmatrix} \mathbf{q}_{\pi1} & \cdots & \mathbf{q}_{\pi6} \end{bmatrix} \begin{bmatrix} \lambda_{\pi1} & & \\ & \ddots & \\ & & \lambda_{\pi6} \end{bmatrix} \begin{bmatrix} \mathbf{q}_{\pi1}^T \\ \vdots \\ \mathbf{q}_{\pi6}^T \end{bmatrix} \tag{3}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_6$ are the eigenvalues of Ψ_π, and $\mathbf{q}_i$ are the corresponding eigenvectors, of which the first three elements are the translation components, and the last three elements are the rotation components. The eigenvector $\mathbf{q}_1$ corresponding to the largest eigenvalue represents the transformation of maximum constraint. Perturbing the plane parameters by the transformation of the direction $\mathbf{q}_1$ will result in the largest possible change in from among all possible transformation perturbations.

4.1.2 Degenerate Cases

Define the matrix $\mathbf{H}$ and compute its SVD decomposition as

$$\mathbf{H} = \sum_{i=1}^{N} {}^r\mathbf{n}_i {}^c\mathbf{n}_i^T = \mathbf{U}_\pi \mathbf{\Lambda}_\pi \mathbf{V}_\pi^T = \lambda_{\pi,1}\mathbf{u}_{\pi,1}\mathbf{v}_{\pi,1}^T + \lambda_{\pi,2}\mathbf{u}_{\pi,2}\mathbf{v}_{\pi,2}^T + \lambda_{\pi,3}\mathbf{u}_{\pi,3}\mathbf{v}_{\pi,3}^T \tag{4}$$

where the singular values $\lambda_{\pi,1}, \lambda_{\pi,2}, \lambda_{\pi,3}$ satisfy $\lambda_{\pi,1} \geq \lambda_{\pi,2} \geq \lambda_{\pi,3}$.

(1) Assuming that $\lambda_{\pi,3} = 0$, ${}^c\mathbf{n}_i^T \mathbf{v}_{\pi,3} = 0$ holds true for all $i = 1, \cdots, N$. For a small camera motion $\delta\mathbf{w} = [\mu\mathbf{v}_{\pi,3}^T, \mathbf{0}^T]^T$ in the direction of $\mathbf{v}_{\pi,3}$, the variation of the cost function $\delta J_\pi(\delta\mathbf{w}) = \delta\mathbf{w}^T \mathbf{\Psi}_\pi \delta\mathbf{w}$ caused by $\delta\mathbf{w}$ is always zero. That is to say, the perturbation in the direction of $\mathbf{v}_{\pi,3}$ will cause no change of the cost function.

(2) Similarly, when $\lambda_{\pi,1} = \lambda_{\pi,2} = 0$, for all $i = 1, \cdots, N$ ${}^c\mathbf{n}_i$ satisfies ${}^c\mathbf{n}_i^T \mathbf{v}_{\pi,2} = 0$, ${}^c\mathbf{n}_i^T \mathbf{v}_{\pi,3} = 0$ and ${}^c\mathbf{n}_i \times \mathbf{v}_{\pi,1} = 0$. In this case, for a small camera motion $\delta\mathbf{w} = [\mu_1\mathbf{v}_{\pi,2}^T + \mu_2\mathbf{v}_{\pi,3}^T, \mu_3\mathbf{v}_{\pi,1}^T]^T$, $\delta J_\pi(\delta\mathbf{w}) = 0$.

4.2 Point based Tracking

4.2.1 Points' Constraints on Camera Motion

Suppose that all the occluding edge points are directly used in the pose estimation.

$$\begin{aligned} J_p(\mathbf{w}) &= \sum_{k=1}^{N_p} J_{p,k} = \sum_{k=1}^{N_p} \mathbf{e}_k^T \mathbf{e}_k \\ &= \sum_{k=1}^{N_p} \left({}^c\mathbf{p}'_{o,k} - R_{cr}\left({}^r\mathbf{p}'_{o,k}\right) - \mathbf{t}'_{cr}\right)^T \left({}^c\mathbf{p}'_{o,k} - R_{cr}\left({}^r\mathbf{p}'_{o,k}\right) - \mathbf{t}'_{cr}\right)\end{aligned} \tag{5}$$

where all the point coordinates are referred to the centroid.

$$\begin{aligned} &\mathbf{t}'_{cr} = \mathbf{t}_{cr} - {}^c\bar{\mathbf{p}}_o + R_{cr}({}^r\bar{\mathbf{p}}_o) \\ &{}^c\mathbf{p}'_{o,k} = {}^c\mathbf{p}_{o,k} - {}^c\bar{\mathbf{p}}_o,\ {}^r\mathbf{p}'_{o,k} = {}^r\mathbf{p}_{o,k} - {}^r\bar{\mathbf{p}}_o \\ &{}^c\bar{\mathbf{p}}_o = \frac{1}{N_p}\sum_{k=1}^{N_p} {}^c\mathbf{p}_{o,k},\ {}^r\bar{\mathbf{p}}_o = \frac{1}{N_p}\sum_{k=1}^{N_p} {}^r\mathbf{p}_{o,k}\end{aligned} \tag{6}$$

The scatter matrix $\mathbf{\Psi}_p$ is defined likewise.

$$\mathbf{\Psi}_p = \sum_{k=1}^{N_p} \left(\frac{\partial J_{p,k}}{\partial \mathbf{w}}\right)\left(\frac{\partial J_{p,k}}{\partial \mathbf{w}}\right)^T \tag{7}$$

where

$$\frac{\partial J_{p,k}}{\partial \mathbf{w}} = -2 \begin{bmatrix} \mathbf{I}_{3\times 3} \\ {}^r\hat{\mathbf{p}}'_{o,k} \end{bmatrix} \left({}^c\mathbf{p}'_{o,k} - R_{cr}({}^r\mathbf{p}'_{o,k}) - \mathbf{t}_{cr}\right) \tag{8}$$

4.2.2 Degenerate Cases

(1) It can be seen from $\mathbf{\Psi}_p$ that if $\exists \mathbf{v}$ such that ${}^r\mathbf{p}'_{o,k} \times \mathbf{v} = 0$ for all $k = 1, \cdots, N_p$, i.e., the measured points are collinear (at least two points), $\delta J_p(\delta\mathbf{w}) = \delta\mathbf{w}^T \mathbf{\Psi}_p \delta\mathbf{w} = 0$ for $\delta\mathbf{w} = [\mathbf{0}^T, \mu\mathbf{v}^T]^T$.

(2) If there is only one measured point, ${}^r\mathbf{p}'_{o,k} = \mathbf{0}$. In this case, for any rotation $\omega \in \mathbb{R}^3$, $\delta J_p(\delta\mathbf{w}) = \delta\mathbf{w}^T \mathbf{\Psi}_p \delta\mathbf{w} = 0$ for $\delta\mathbf{w} = [\mathbf{0}^T, \omega^T]^T$.

4.3 Plane and Point based Tracking

The overall objective function is

$$J(\mathbf{w}) = J_\pi(\mathbf{w}) + J_p(\mathbf{w}) \tag{9}$$

4.3.1 Point Weighting

For each occluding edge point ${}^c\mathbf{p}_{o,k}$, a weight $\alpha_{p,k}$ is assigned to the corresponding component in the cost function

$$J_p(\mathbf{w}) = \sum_{k=1}^{N_p} \alpha_{p,k} J_{p,k} \tag{10}$$

$$\alpha_{p,k} = \frac{\left(\frac{\partial J_{p,k}}{\partial \mathbf{w}}\right)^T \mathbf{q}_{\pi,j}}{\left|\frac{\partial J_{p,k}}{\partial \mathbf{w}}\right| \sqrt{\frac{\lambda_{\pi,j}}{\bar{\lambda}_\pi}}} \tag{11}$$

$$j = \arg\max_j \frac{\left(\frac{\partial J_{p,k}}{\partial \mathbf{w}}\right)^T \mathbf{q}_{\pi,j}}{\left|\frac{\partial J_{p,k}}{\partial \mathbf{w}}\right|} \tag{12}$$

$$\bar{\lambda}_\pi = \frac{1}{6} \sum_{t=1}^{6} \lambda_{\pi,t} \tag{13}$$

4.3.2 Degenerate Cases

Suppose that the measurements include both planes and points. Combining the degenerate cases stated in 1.1.2 and 1.2.2 yields the following cases.

(1) $\lambda_{\pi,1} = \lambda_{\pi,2} = 0$ and ${}^r\mathbf{p}'_{o,k} \times \mathbf{v} = 0$ for all $k = 1, \cdots, N_p$ hold true simultaneously, and $\mathbf{v}$ and $\mathbf{v}_{\pi,1}$ happen to be of the same direction, i.e., $\mathbf{v} \times \mathbf{v}_{\pi,1} = 0$. In this case, the camera motion along $\delta\mathbf{w} = [\mathbf{0}^T, \mu\mathbf{v}_{\pi,1}^T]^T$ cannot be constrained.

(2) If $\lambda_{\pi,1} = \lambda_{\pi,2} = 0$ and there is only one measured point, the camera motion along $\delta\mathbf{w} = [\mathbf{0}^T, \mu\mathbf{v}_{\pi,1}^T]^T$ cannot be constrained.

5 Plane and Point based Camera Tracking with Unknown Correspondences

The plane segmentation method proposed in [1] and the edge detection method proposed in [2] are used to detect the planar segments and edge points in an RGB-D scan.

A detected planar segment $\mathcal{P}_i, i = 1, \cdots, N$ has the following attributes.

- plane parameters $\pi_i = [\mathbf{n}_i^T, d_i]^T$
- number of points $N_{\pi,i}$
- centroid $\mathbf{p}_{\pi,i}$
- covariance $\mathbf{C}_{\pi,i}$

Suppose that $\{{}^r\mathcal{P}_i\}_{i=1,\cdots,N_r}$ and $\{{}^c\mathcal{P}_j\}_{j=1,\cdots,N_c}$ are planar segments extracted from reference and current frame, respectively. Each planar segment $\mathcal{P}_i$ is modeled as a Gaussian distribution $\mathcal{N}(\mathbf{p}_{\pi,i}, \mathbf{C}_{\pi,i})$. $\{{}^r\mathbf{p}_{o,k}\}_{k=1,\cdots,N_{pr}}$ and $\{{}^c\mathbf{p}_{o,l}\}_{l=1,\cdots,N_{pc}}$ are occluding edge points extracted from reference and current frame, respectively. The planar segments and occluding edge points extracted from two successive frames are used simultaneously in a ICP framework. In each iteration of ICP, the correspondences between planar segments are assigned by checking the Bhattacharyya distance between the Gaussian distribution that models the planar segments.

$$D_{Bha}(\mathcal{P}_i, \mathcal{P}_j) = \frac{1}{8} \left(\mathbf{p}_{\pi,i} - \mathbf{p}_{\pi,j}\right)^T \mathbf{C}_\pi^{-1} \left(\mathbf{p}_{\pi,i} - \mathbf{p}_{\pi,j}\right) + \frac{1}{2} \ln\left(\frac{|\mathbf{C}_\pi|}{\sqrt{|\mathbf{C}_{\pi,i}||\mathbf{C}_{\pi,j}|}}\right) \tag{14}$$

Table 1:

	Plane-Point ICP	STING
Fr1/xyz	0.0123	**0.0114**
Fr1/room	**0.0817**	0.0832
Fr3/cabinet	0.0382	**0.0326**

where $\mathbf{C}_\pi = \frac{1}{2}(\mathbf{C}_{\pi,i} + \mathbf{C}_{\pi,j})$. The correspondences between points are assigned by checking the Euclidean distance between the point coordinates.

$$D_{Euc}(\mathbf{p}_{o,k}, \mathbf{p}_{o,l}) = \|\mathbf{p}_{o,k} - \mathbf{p}_{o,l}\|_2 \tag{15}$$

The whole ICP framework is presented in Algorithm 1.

Algorithm 1 Planar Segment and Occluding Edge Point based ICP

Inputs:

Planar segments extracted from two frames $\{^r\mathcal{P}_i\}_{i=1,\cdots,N_r}$ and $\{^c\mathcal{P}_j\}_{j=1,\cdots,N_c}$.

Occluding edge points extracted from two frames $\{^r\mathbf{p}_{o,k}\}_{k=1,\cdots,N_{pr}}$ and $\{^c\mathbf{p}_{o,l}\}_{l=1,\cdots,N_{pc}}$.

Outputs:

Transformation between two frames $T_{cr} = \mathbf{R}_{cr}, \mathbf{t}_{cr}$.

6 Plane Fusion

A plane $\mathcal{P}_i$

- plane parameters $\pi_i = [\mathbf{n}_i^T, d_i]^T$
- number of points $N_{\pi,i}$
- centroid $\mathbf{p}_{\pi,i}$
- covariance $\mathbf{C}_{\pi,i}$
- curvature $\rho_{\pi,i}$ [1]
- shadow points $\left\{\mathbf{p}_{s,k}^i, k = 1, \cdots, N_p^i\right\}$

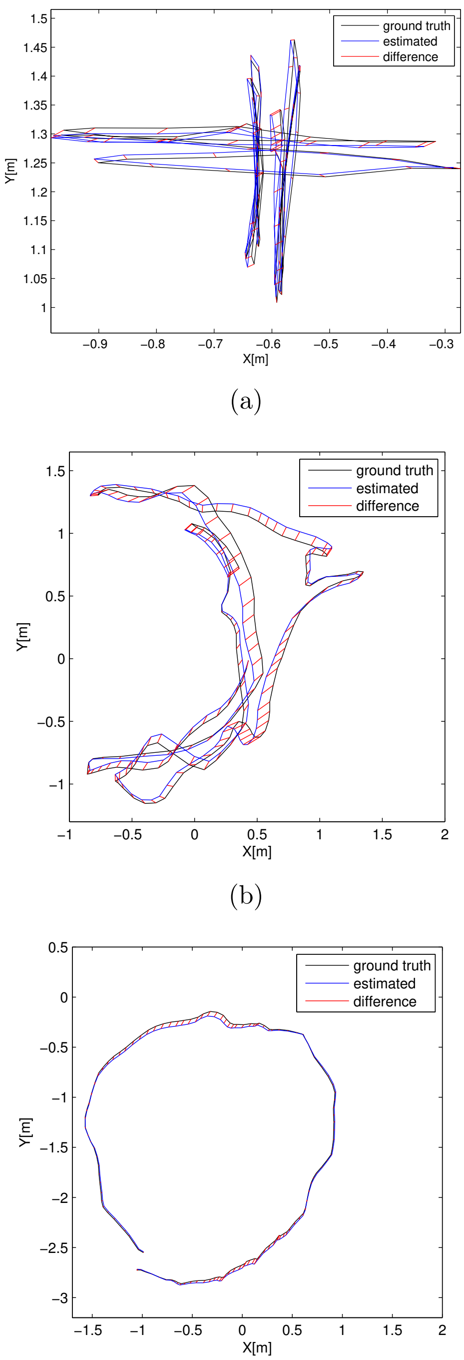
7 Least Primitives for Pose Estimation

7.1 Three Planes

Three corresponding non-parallel planes $\{^r\pi_i, {}^c\pi_i, \}_{i=1,2,3}$. Let

$$\begin{aligned} {}^r\mathbf{M}_1 &= \begin{bmatrix} {}^r\mathbf{n}_1 & {}^r\mathbf{n}_2 & {}^r\mathbf{n}_3 \end{bmatrix} \\ {}^c\mathbf{M}_1 &= \begin{bmatrix} {}^c\mathbf{n}_1 & {}^c\mathbf{n}_2 & {}^c\mathbf{n}_3 \end{bmatrix} \\ \mathbf{d} &= \begin{bmatrix} {}^cd_1 - {}^rd_1 \\ {}^cd_2 - {}^rd_2 \\ {}^cd_3 - {}^rd_3 \end{bmatrix} \end{aligned} \tag{16}$$

[1]Note that the curvature here is just an indication that tells how a surface deviates from being a flat plane, rather than the strictly defined curvature.

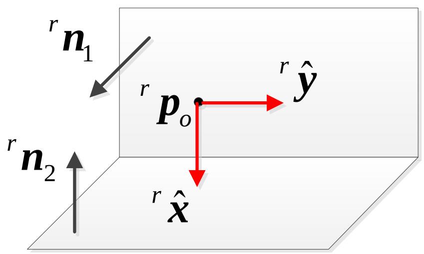


(a)

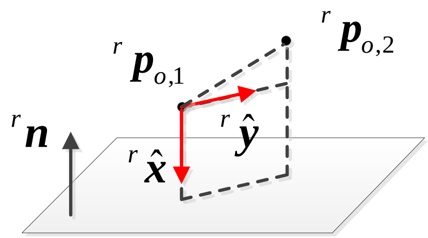
(b)

(c)

Figure 1: (a)Fr1/xyz; (b)Fr1/room; (c)Fr3/cabinet;



(a)



(b)

Figure 2:

The rotation $\mathbf{R}_{cr}$ and translation $\mathbf{t}_{cr}$ can be computed as

$$\begin{aligned} \mathbf{R}_{cr} &= {}^c\mathbf{M}_1 {}^r\mathbf{M}_1^{-1} \\ \mathbf{t}_{cr} &= {}^c\mathbf{M}_1^{-T}\mathbf{d} \end{aligned} \tag{17}$$

7.2 Two Planes and One Point

Two corresponding non-parallel planes $\{{}^r\pi_i, {}^c\pi_i,\}_{i=1,2}$ and one corresponding point $\{{}^r\mathbf{p}_o, {}^c\mathbf{p}_o\}$. Construct three axes ${}^r\hat{\mathbf{x}}, {}^r\hat{\mathbf{y}}, {}^r\hat{\mathbf{z}}$ in reference coordinate system and locate the origin at ${}^r\mathbf{p}_o$.

$$\begin{aligned} {}^r\hat{\mathbf{x}} &= -{}^r\mathbf{n}_2 \\ {}^r\hat{\mathbf{y}} &= {}^r\mathbf{n}_1 \times {}^r\mathbf{n}_2 \\ {}^r\hat{\mathbf{z}} &= {}^r\hat{\mathbf{x}} \times {}^r\hat{\mathbf{y}} \end{aligned} \tag{18}$$

The axes and the origin in current frame is constructed likewise.

$$\begin{aligned} {}^c\hat{\mathbf{x}} &= -{}^c\mathbf{n}_2 \\ {}^c\hat{\mathbf{y}} &= {}^c\mathbf{n}_1 \times {}^c\mathbf{n}_2 \\ {}^c\hat{\mathbf{z}} &= {}^c\hat{\mathbf{x}} \times {}^c\hat{\mathbf{y}} \end{aligned} \tag{19}$$

Let

$$\begin{aligned} {}^r\mathbf{M}_2 &= \begin{bmatrix} {}^r\hat{\mathbf{x}} & {}^r\hat{\mathbf{y}} & {}^r\hat{\mathbf{z}} \end{bmatrix} \\ {}^c\mathbf{M}_2 &= \begin{bmatrix} {}^c\hat{\mathbf{x}} & {}^c\hat{\mathbf{y}} & {}^c\hat{\mathbf{z}} \end{bmatrix} \end{aligned} \tag{20}$$

Then the rotation $\mathbf{R}_{cr}$ and translation $\mathbf{t}_{cr}$ can be computed as

$$\begin{aligned} \mathbf{R}_{cr} &= {}^c\mathbf{M}_2 {}^r\mathbf{M}_2^T \\ \mathbf{t}_{cr} &= {}^c\mathbf{p}_o - \mathbf{R}_{cr} {}^r\mathbf{p}_o \end{aligned} \tag{21}$$

7.3 One Plane and Two Points

One corresponding plane $\{{}^r\pi, {}^c\pi,\}$ and two different corresponding points $\{{}^r\mathbf{p}_{o,j}, {}^c\mathbf{p}_{o,j}\}_{j=1,2}$ satisfying $({}^r\mathbf{p}_{o,1} - {}^r\mathbf{p}_{o,2}) \times {}^r\mathbf{n} \neq 0$ and $({}^c\mathbf{p}_{o,1} - {}^c\mathbf{p}_{o,2}) \times {}^c\mathbf{n} \neq 0$. Construct three axes ${}^r\hat{\mathbf{x}}, {}^r\hat{\mathbf{y}}, {}^r\hat{\mathbf{z}}$ in reference coordinate system and locate the origin at ${}^r\mathbf{p}_{o,1}$.

$$\begin{aligned} {}^r\hat{\mathbf{x}} &= -{}^r\mathbf{n} \\ {}^r\mathbf{y} &= ({}^r\mathbf{p}_{o,2} - {}^r\mathbf{p}_{o,1}) - \left(({}^r\mathbf{p}_{o,2} - {}^r\mathbf{p}_{o,1})^T \, {}^r\mathbf{n} \right) {}^r\mathbf{n} \\ {}^r\hat{\mathbf{y}} &= \frac{{}^r\mathbf{y}}{\|{}^r\mathbf{y}\|} \\ {}^r\hat{\mathbf{z}} &= {}^r\hat{\mathbf{x}} \times {}^r\hat{\mathbf{y}} \end{aligned} \tag{22}$$

The axes and the origin in current frame is constructed likewise.

$$
\begin{aligned}
{}^c\hat{\mathbf{x}} &= -{}^c\mathbf{n} \\
{}^c\mathbf{y} &= ({}^c\mathbf{p}_{o,2} - {}^c\mathbf{p}_{o,1}) - \left(({}^c\mathbf{p}_{o,2} - {}^c\mathbf{p}_{o,1})^T\,{}^c\mathbf{n}\right){}^c\mathbf{n} \\
{}^c\hat{\mathbf{y}} &= \frac{{}^c\mathbf{y}}{\|{}^c\mathbf{y}\|} \\
{}^c\hat{\mathbf{z}} &= {}^c\hat{\mathbf{x}} \times {}^c\hat{\mathbf{y}}
\end{aligned} \tag{23}
$$

Let

$$
\begin{aligned}
{}^r\mathbf{M}_3 &= \begin{bmatrix} {}^r\hat{\mathbf{x}} & {}^r\hat{\mathbf{y}} & {}^r\hat{\mathbf{z}} \end{bmatrix} \\
{}^c\mathbf{M}_3 &= \begin{bmatrix} {}^c\hat{\mathbf{x}} & {}^c\hat{\mathbf{y}} & {}^c\hat{\mathbf{z}} \end{bmatrix}
\end{aligned} \tag{24}
$$

Then the rotation $\mathbf{R}_{cr}$ and translation $\mathbf{t}_{cr}$ can be computed as

$$
\begin{aligned}
\mathbf{R}_{cr} &= {}^c\mathbf{M}_3\,{}^r\mathbf{M}_3^T \\
\mathbf{t}_{cr} &= {}^c\mathbf{p}_{o,1} - \mathbf{R}_{cr}\,{}^r\mathbf{p}_{o,1}
\end{aligned} \tag{25}
$$

References

[1] Efficient Organized Point Cloud Segmentation with Connected Components, SPME, 2013.

[2] RGB-D Edge Detection and Edge-based Registration, IROS, 2013.