Assignment 4

Andrew ID: sajaved

Name: Syed Ashar Javed

Q2

Part a

The point-to-point ICP minimizes the distance between the transformed points (using the estimated pose) and the original point cloud which we are aligning with. The point-to-plane ICP however minimizes the distance of the transformed points to the plane on which the original point clouds lie.

Point-to-Point ICP:

$$E = \min_{T} \|Tv - \hat{v}\|$$

Point-to-Plane ICP:

$$E = \min_{T} \left\|(Tv - \hat{v})^T.N\right\|$$

where T is the transformation pose, v is the 3d input point, $\hat{v}$ is the closest point from the existing point cloud and N is the vector normal plane to the surface of the point. cloud.

Part b

The derivation for rigid body transformation approximation is given below:

$$\tilde{T}^z_{g,k}\dot{V}_k(u) = \tilde{R}^z\tilde{V}^g_k(u) + \tilde{t}^z$$

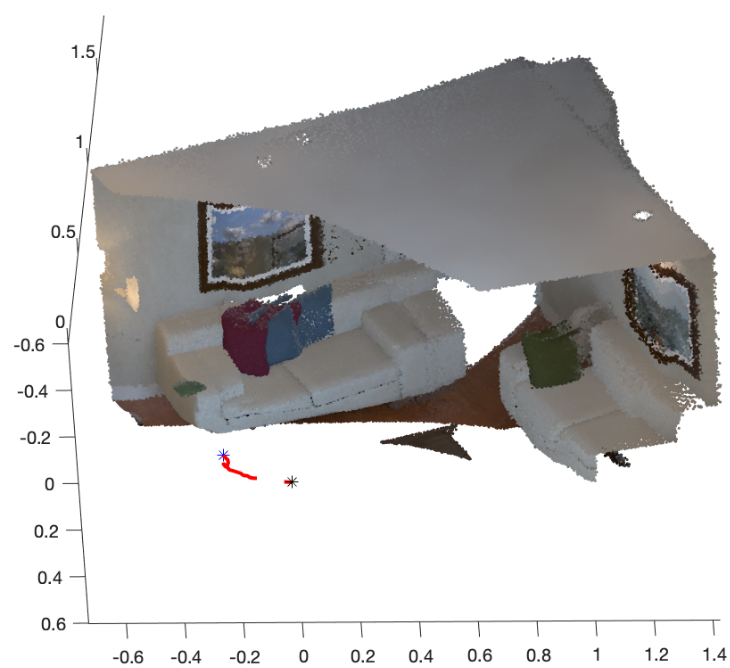
Expanding the rotation matrix and approximating the terms with a small angle approximation, the trigonometric functions are replaced by just their angles as done below and the terms are rearranged to get the matrix G and V:

$$
\begin{aligned}
\tilde{T}^z_{g,k}\dot{V}_k(u) &= \begin{bmatrix} 1 & \alpha & -\gamma \\ -\alpha & 1 & \beta \\ \gamma & -\beta & 1 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \\
&= \begin{bmatrix} v_x & \alpha v_y & -\gamma v_z \\ -\alpha v_x & v_y & \beta v_z \\ \gamma v_x & -\beta v_y & v_z \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \\
&= \begin{bmatrix} 0.\beta & -\gamma w & \alpha v & t_x & 0.t_y & 0.t_z \\ \beta w & -0.\gamma & \alpha u & 0.t_x & t_y & 0.t_z \\ -\beta v & +\gamma u & 0.\alpha & 0.t_x & 0.t_y & t_z \end{bmatrix} + \begin{bmatrix} u \\ v \\ w \end{bmatrix} \\
&= \left[\begin{array}{ccc|ccc} 0 & -v_z & v_y & 1 & 0 & 0 \\ v_z & 0 & -v_x & 0 & 1 & 0 \\ -v_y & v_x & 0 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} \beta \\ \gamma \\ \alpha \\ t_x \\ t_y \\ t_z \end{bmatrix} + \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} \\
&= \mathbf{G}(\mathbf{u})\mathbf{x} + \tilde{\mathbf{V}}^g_k(\mathbf{u})
\end{aligned}
$$

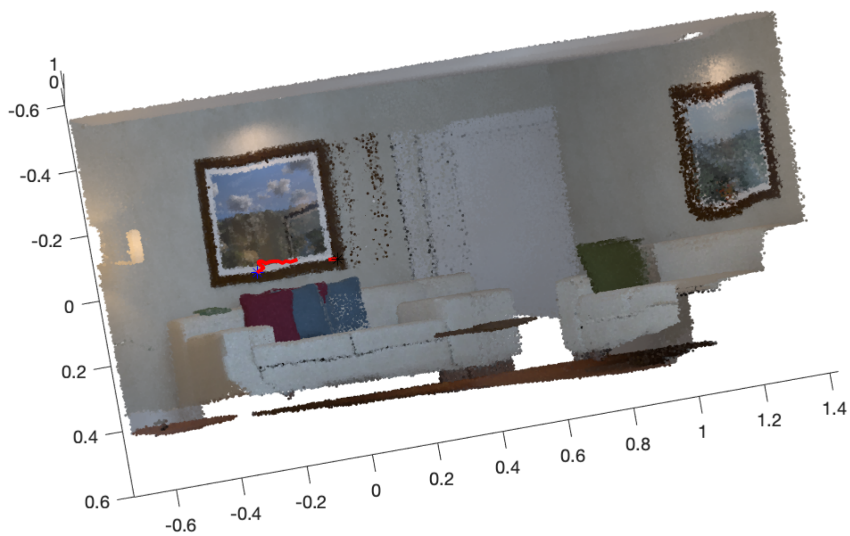
where $\tilde{\mathbf{V}}^g_k(\mathbf{u}) = u = (v_x, v_y, v_z)$, $G(u) \in R^{3\times 6}$ is the first matrix, $\mathbf{x} = (\beta, \gamma, \alpha, t_x, t_y, t_z)$ is the set of transformation parameters.

Part d

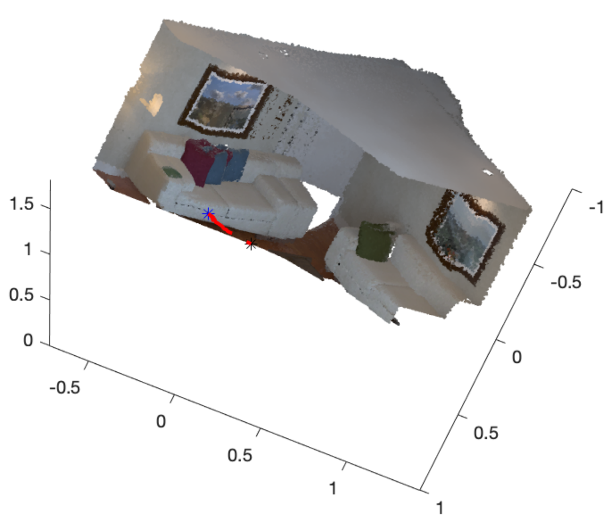
The below results are from the ICP code by putting $is\_debug\_icp = 1$.



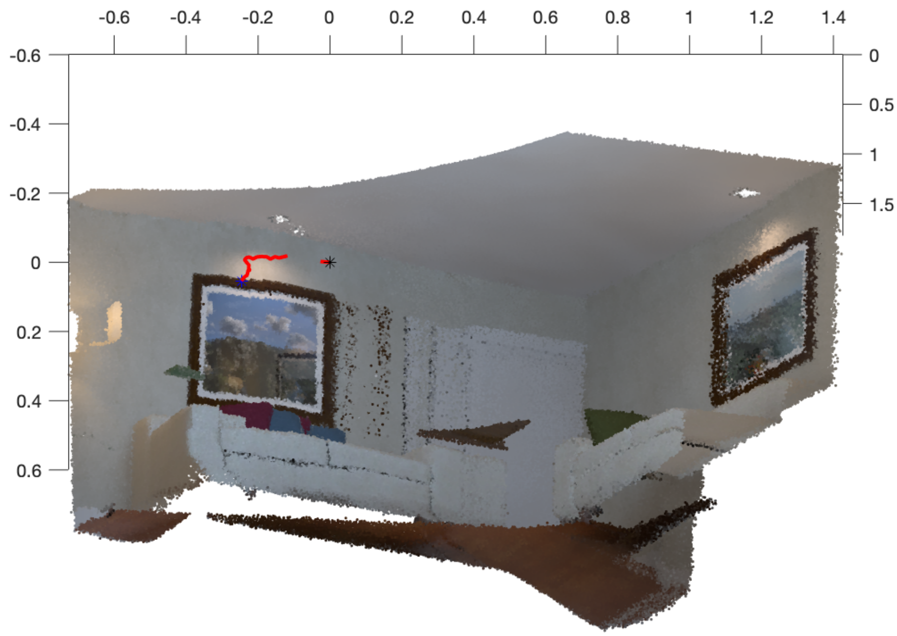
3D point clouds: View 1



3D point clouds: View 2



3D point clouds: View 3



3D point clouds: View 4

The summary from running the ICP are:

```
Number of points in the final model = 1072357
Compression ratio = 13.43 %
Total time spent = 9.53 sec
```

ICP Summary

Q3

Part a

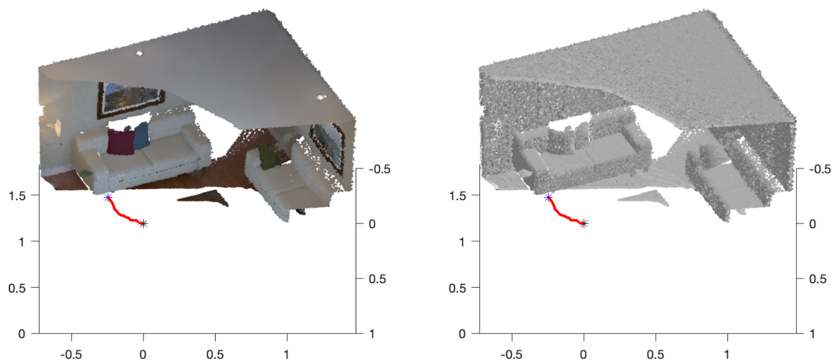
Volumetric representations are memory intensive and are less flexible since they often require fixing the resolution beforehand for efficiency reasons. This also means that the whole volume might need to be updated at times like loop closures which adds to the inflexibility. These limitations also impose scalability constraints to volumetric representations. Finally, as mentioned in the point-based fusion paper of Keller et al., point based methods have less overhead from the transitions required between different representations in volumetric settings.

Part b

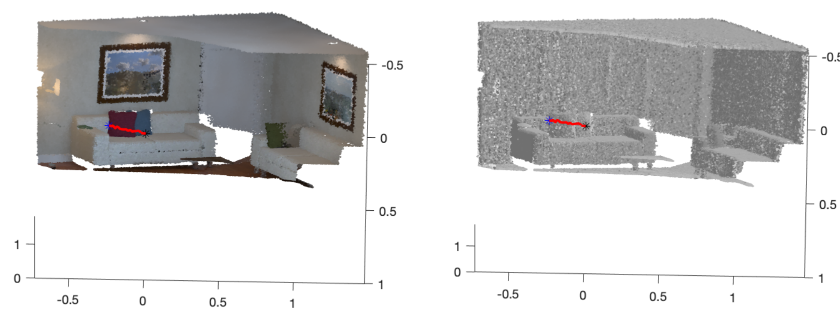
Since the normal n is perpendicular to the vector p which is transformed as $p' = Rp + t$, the normal vector can be written in terms of the rotation component of the transformation, ie $n' = Rn$.

Part d

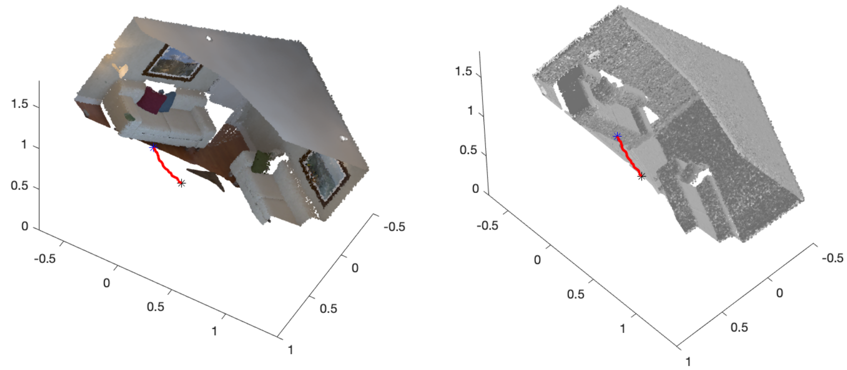
The results below are from the ICP + Fusion code.



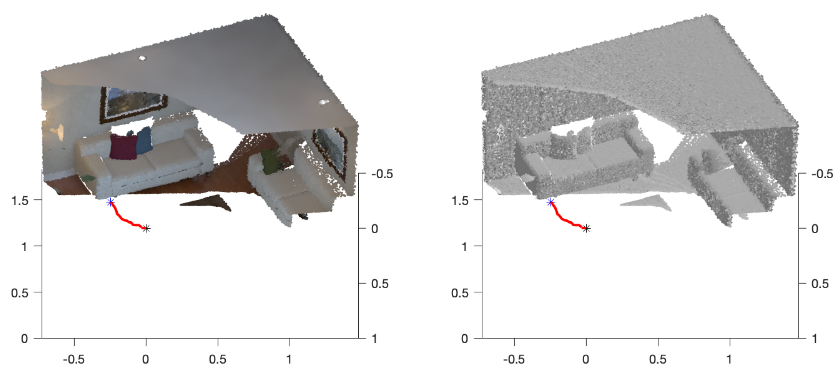
3D point clouds: View 1



3D point clouds: View 2



3D point clouds: View 3



3D point clouds: View 4

The summary from running the ICP are:

```
Number of points in the final model = 592533
Compression ratio = 7.42 %
Total time spent = 13.35 sec
```

ICP + Fusion Summary

Q4

Part a

Yes, it does. Using the current frame leads to propagation of errors to the next frame, however in the fusion model, the errors are not propagated into the next frame and we get better registration.

Part b

The inliers and the RMSE are the metrics used for the ICP registration. According to the metrics, the more inliers you have, the better is the registration. Also, lower is the RMSE, the better is the result.

The compression ratio is the metric used for point based fusion. According to this metric, a lower compression ratio is better since the model is able to represent the model with less number of points.

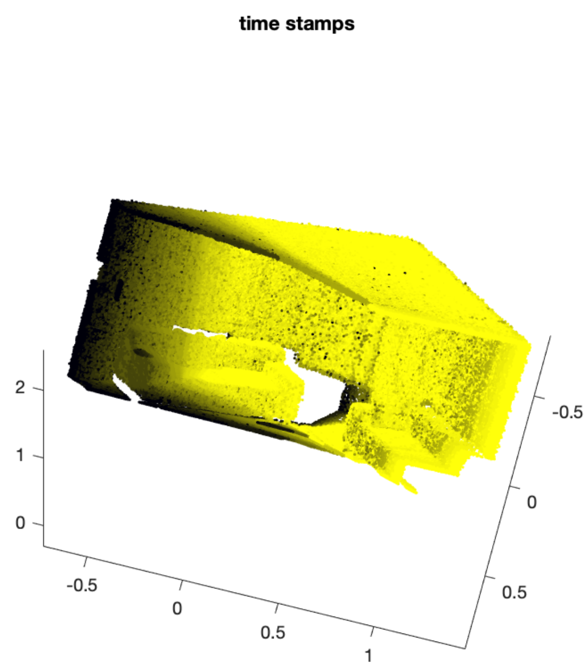
Part c

The confidence counts on the left side are brighter than the ones on the right side signifying that the values on the left are much larger than the values on the right. This is particularly true because the camera spends more time and has more frames for the left side than the right side (camera moving from left to right).

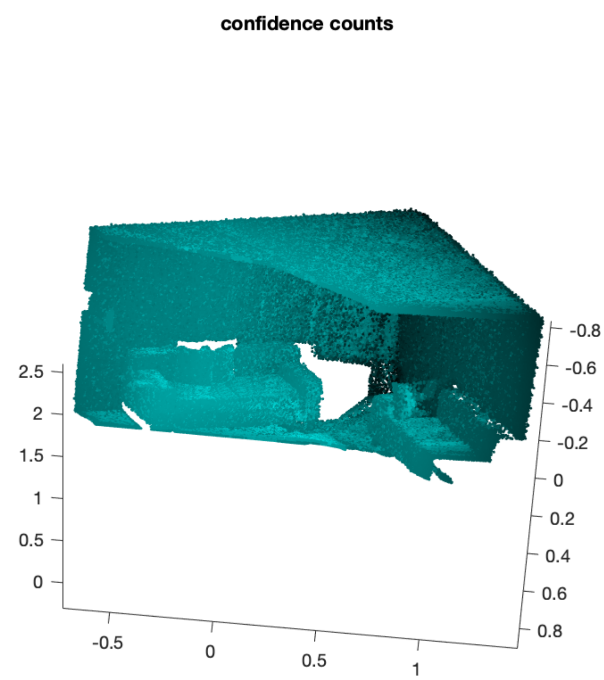
Normals seem to be mapping correctly as can be seen in the images below. For example the normals for floor and ceiling seem to be pointing in the opposite and correct direction.

The time stamps for the left side of the reconstruction has lower values than the time stamps for the right side, since the camera moves from left to right. Hence the left side is darker and the right side is bright.

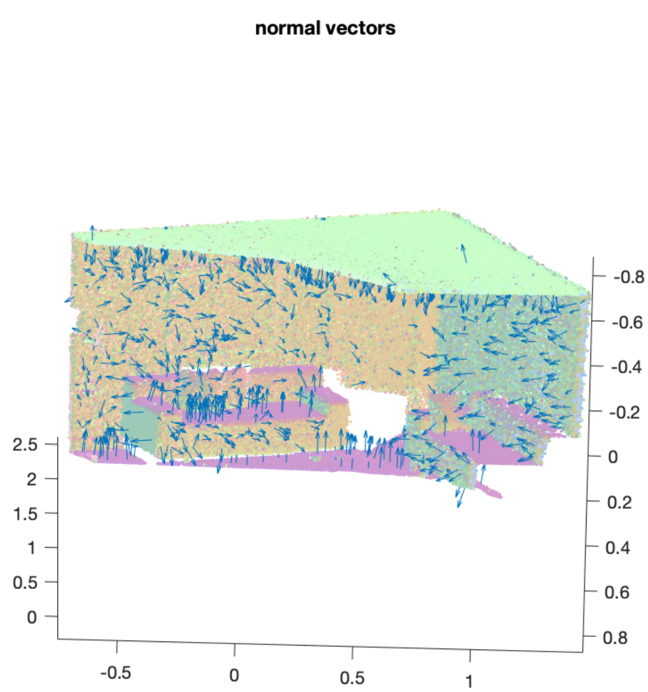
The timestamps, confidence counts and normals are shown below:



Timestamp map



Confidence count map



Normal map