Lukas Riethmüller
2025403370

Setting up

1 Preperation (Point Clouds)

Generating the point cloud from the files in fruit.zip proves to be a very long task. The pipeline was as follows: first the feature extraction happens (this took 13 minutes in google colab), i.e. the program finds the interesting features in the individual images. Then the feature matching begins (google colab was too slow to finish this part and I was not able to follow through), where features across different images are identified, linking their matches. There should also be a last step (I could not get to with google colab), where the program attempts to reconstruct the original object from the now matched features in the form of a mesh.
The next task is to visualize an already processed mesh with open3D, the results are seen in this figure:

2 Building Gaussian Primitives

A general 3D Gaussian is defined thorugh its mean $\boldsymbol{\mu}$ and covariance Σ. If we obtain an ellipsoid by applying a linear map A to an isotropic unit Gaussian, i.e. $\mathbf{y} = \boldsymbol{\mu} + A\mathbf{x}$ with $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, I)$, then the covariance of $\mathbf{y}$ is

$$\Sigma = AA^T.$$

Let $S = \operatorname{diag}(s_x, s_y, s_z)$ be a scaling matrix and let $R \in \mathrm{SO}(3)$ be a rotation matrix. Suppose $A = R\,S$, then we get

$$\Sigma \;=\; (RS)(RS)^T \;=\; RSS^TR^T \;=\; R\,S^2\,R^T,$$

where $S^2 = \operatorname{diag}(s_x^2, s_y^2, s_z^2)$. If we assume general gaussian with coveriance Σ_0, then

$$\Sigma \;=\; RS\Sigma_0S^TR^T.$$

A viewing transformation $\mathbf{T}_v$ is a 4x4 matrix a rotation part $A_v \in \mathbb{R}^{3\times3}$ and translation $\mathbf{t}_v \in \mathbb{R}^3$,

$$\mathbf{T}_v = \begin{pmatrix} A_v & \mathbf{t}_v \\ \mathbf{0}^T & 1 \end{pmatrix}, \qquad \mathbf{y} = A_v\mathbf{x} + \mathbf{t}_v.$$

Then

$$\mathcal{N}(\boldsymbol{\mu}, \Sigma) \xrightarrow{\mathbf{T}_v} \mathcal{N}\big(A_v\boldsymbol{\mu} + \mathbf{t}_v,\; A_v\,\Sigma\,A_v^T\big).$$

The translation only shifts the mean, while the rotation effects both covariance and mean.

For a nonlinear projection transformation $\pi(\mathbf{x})$, linearize it at the mean $\boldsymbol{\mu}$ with Jacobian

$J = \frac{\partial \pi}{\partial \mathbf{x}}\big|_{\boldsymbol{\mu}}$

$$\pi(x) \approx J(\mu)(x-\mu) + \pi(\mu).$$

This linearly transforms our covariance as follows:

$$\Sigma_{\text{proj}} \approx J\,\Sigma\,J^T.$$

We now want to find a square bounding-box side length that cover 99% of the confidence region. Its probability level for a gaussian is defined through χ^2 quantile

$$(x-\mu)^T\Sigma_{\text{proj}}^{-1}(x-\mu) = \chi^2_{2;0.99}, \qquad \chi^2_{2;0.99} \approx 9.21034037197618, \qquad k := \sqrt{\chi^2_{2;0.99}} \approx 3.034854$$

Assuming our Covariance is diagonalized (else do so), we can express the semi-axis lengths as follows:

$$\max_{(x-\mu)^T\Sigma^{-1}(x-\mu)\le k^2} |x_x-\mu_x| = k\sqrt{\Sigma_{11}}, \qquad \max_{(x-\mu)^T\Sigma^{-1}(x-\mu)\le k^2} |x_y-\mu_y| = k\sqrt{\Sigma_{22}}$$

To have a confident cover use the largest semi axis for the squared cover:

$$\ell_{\text{square}} \;=\; 2k\;\max\!\left(\sqrt{\Sigma_{11}}, \sqrt{\Sigma_{22}}\right) \;\approx\; 2\cdot 3.034854\;\max(\sigma_x,\sigma_y) \;\approx\; 6.069708\;\max(\sigma_x,\sigma_y),$$

where $\sigma_x = \sqrt{\Sigma_{11}}$ and $\sigma_y = \sqrt{\Sigma_{22}}$.