Reconstruction by concatenating multiple stereo vision

- For calibrated cameras, stereo reconstruction possible
- Camera calibration:
  - Intrinsic parameters: by chessboard-based Zhang calibration
  - Extrinsic parameters: by decomposition of essential matrix
- Spatial reconstruction: triangulation
- Results:
  - For each stereo image pair, 3D point clouds obtained.
  - Transformation(translation/rotation) between images computed as well

Concatenating point clouds

- Two point clouds given
  - For stereo reconstruction, coordinate system is usually fixed to the first camera.
- Point clouds have N common points are stecked in vector sets: $\{\mathbf{p_i}\}$ and $\{\mathbf{q_i}\}$, $(i = 1 \dots N)$.
- Similarity transformation between images has o be estimated.

$$\mathbf{q_i} = s\mathbf{R}\mathbf{p_i} + \mathbf{t}$$

  - s: scale
  - $\mathbf{R}$: rotation
  - $\mathbf{t}$: translation

Concatenating stereo reconstructions

- Task: optimal registration to estimate similarity transformation

$$\sum_{i=1}^{N} ||\mathbf{q_i} - s\mathbf{R}\mathbf{p_i} - \mathbf{t}||^2$$

- Proof given in separate document
  - Optimal translation $\mathbf{t}$: difference of centers of gravity
  - Optimal rotation:

$$\mathbf{H} = \sum_{i=1}^{N} \mathbf{q'_i}\mathbf{p'_i}^T$$

$$\mathbf{R} = \mathbf{V}\mathbf{U}^\mathbf{T} \leftarrow \mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^T$$

  - Optimal scale:

$$s = \frac{\sum_{i=1}^{N} \mathbf{q'_i}^\mathbf{T} R\mathbf{p'_i}}{\sum_{i=1}^{N} \mathbf{p'_i}^\mathbf{T}\mathbf{p'_i}}$$