



# Lecture 08 – Pose Estimation I

Nov 11<sup>nd</sup>, 2018

Danping Zou,

Associate Professor

Institute for Sensing and Navigation



上海交通大學  
SHANGHAI JIAO TONG UNIVERSITY

# Outline

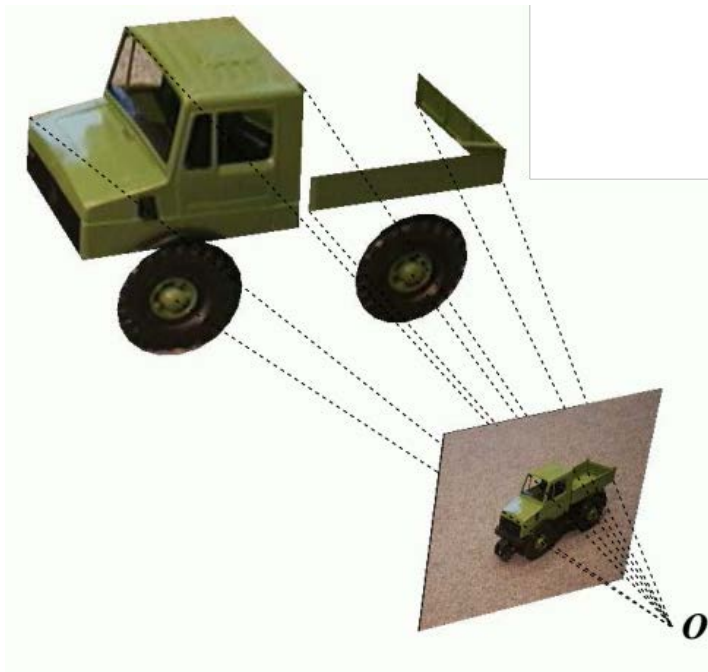


- **About pose estimation**
- **3D-to-3D registration**
  - Rotation only
  - Rotation plus Translation
  - Unknown correspondences – Iterative Closest Point (ICP)
- **3D-to-2D registration (Camera pose estimation)**
  - 3D objects
    - Close-form algorithm - P3P
    - Iterative algorithm - POSIT
    - Iterative algorithm - Nonlinear least squares
  - Planar objects
    - Known patterns – Checkboard box, QR pattern
    - Planar Pictures

# Pose estimation



- Given a 3D model and its projection on the image, we want to get the camera pose with respect to the 3D model.
- This process can also be treated as 3D-2D registration problem

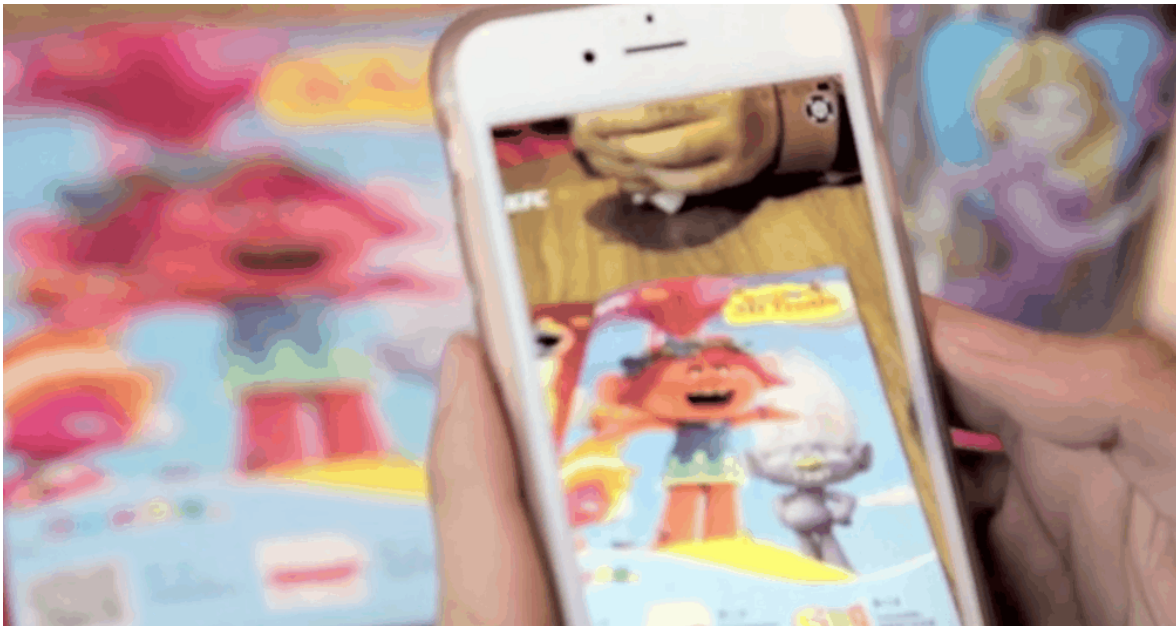


$$X \leftrightarrow x \Rightarrow R, t$$

# Pose estimation



- Pose estimation is a basic problem in augmented reality.
- About Augmented reality

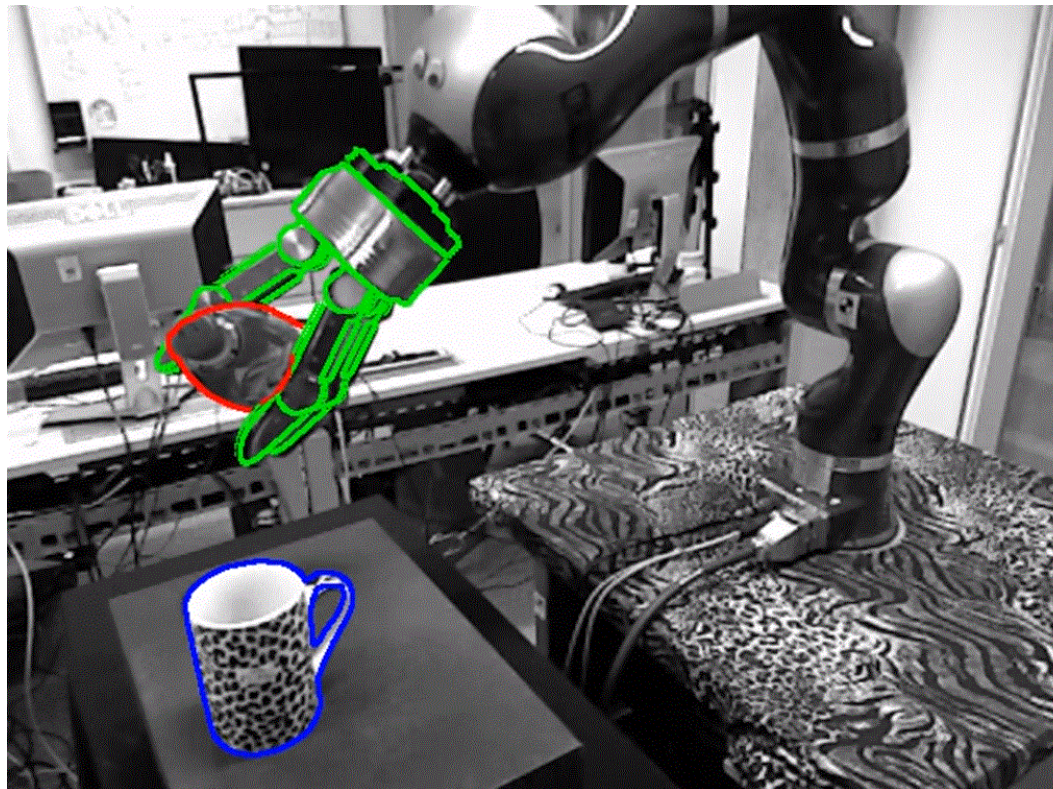




# Pose estimation



- Pose estimation is also critical for grasping and manipulation in robotics.



# Pose estimation




- Pose estimation is also applied to image-based localization
  - use 3D reconstruction method to generate 3D point clouds of the scene
  - extract the feature points from the query image and match them to the 3D points (Pose estimation)



**Query image**

**Feature  
matching**



**Pose  
estimation**



**3D point cloud from internet photos**

# Pose estimation



- We first discuss about 3D-3D registration and then discuss about 2D-2D registration



$$\begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$



$$\mathbf{K}[\mathbf{R} \ \mathbf{t}]$$



# Rotation-only 3D-3D registration



- Rotation-only 3D-to-3D registration (Recall the last homework)
  - The corresponding points are known:

$$\mathbf{x}_i \leftrightarrow \mathbf{y}_i$$

- Only rotation between two point clouds is applied:

$$\mathbf{y}_i = \mathbf{R}\mathbf{x}_i$$

- Our problem is to solve the rotation  $\mathbf{R}$  from the corresponding points



# Rotation-only 3D-3D registration



- In the last homework, we use Gauss-Newton method optimize the object function iteratively,

$$\min_{\mathbf{R}} \sum_{i=1}^n (\mathbf{y}_i - \mathbf{R}\mathbf{x}_i)^2$$

- We start from an initial rotation :  $\mathbf{R} \leftarrow \mathbf{R}_0$
- And iteratively solve the incremental parameter  $\Delta\theta \in \mathbb{R}^{3 \times 1}$

$$\mathbf{R} \leftarrow \mathbf{R} \boxplus \Delta\theta \quad (\mathbf{R} \leftarrow \mathbf{R} \exp([\Delta\theta]_{\times}))$$

# Rotation-only 3D-3D registration



- $\Delta\theta$  is solved by minimizing

$$\sum_{i=1}^n \|\mathbf{y}_i - \mathbf{R} \exp(\Delta\theta^\wedge) \mathbf{x}_i\|^2$$

$$\approx \sum_{i=1}^n \|\mathbf{y}_i - \mathbf{R}(\mathbf{I} + [\Delta\theta]_\times) \mathbf{x}_i\|^2 \quad (\text{First order approximation})$$

$$\sum_{i=1}^n \|\mathbf{y}_i - \mathbf{R} \mathbf{x}_i - \mathbf{R} [\Delta\theta]_\times \mathbf{x}_i\|^2$$

$$\sum_{i=1}^n \|\mathbf{y}_i - \mathbf{R} \mathbf{x}_i + [\mathbf{R} \mathbf{x}_i]_\times \Delta\theta\|^2 \quad (a \times b = [a]_\times b = -b \times a = -[b]_\times a)$$

# Rotation-only 3D-3D registration



- It is a linear least squares problem :

$$\mathbf{J}\Delta\theta = \mathbf{z}$$

$$\begin{bmatrix} -[\mathbf{R}\mathbf{x}_1] \times \\ -[\mathbf{R}\mathbf{x}_1] \times \\ \dots \\ -[\mathbf{R}\mathbf{x}_n] \times \end{bmatrix} \Delta\theta = \begin{bmatrix} \mathbf{y}_1 - \mathbf{R}\mathbf{x}_1 \\ \mathbf{y}_2 - \mathbf{R}\mathbf{x}_2 \\ \dots \\ \mathbf{y}_n - \mathbf{R}\mathbf{x}_n \end{bmatrix}$$

$$\rightarrow \Delta\theta = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{z}$$

- We can find that the algorithm usually converges very fast even if the initial guess is not good.

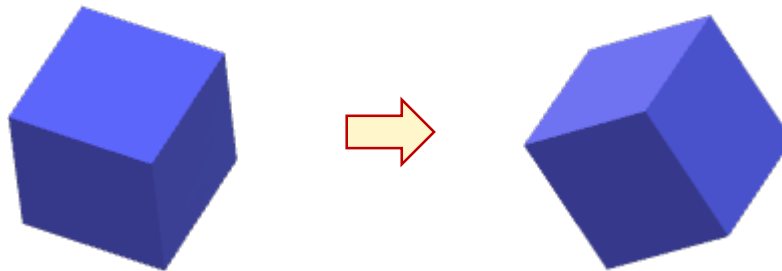
# Rotation-only 3D-3D registration



- In fact, we can solve the rotation in close form.

$$\mathbf{y}_i = \mathbf{R}\mathbf{x}_i$$

$$(i = 1, 2, \dots)$$





# Rotation-only 3D-3D registration



- **First approach** – Direct Linear Transformation (DLT)

$$\left. \begin{aligned} y_1 &= \mathbf{R}x_1 \\ y_2 &= \mathbf{R}x_2 \\ &\dots \\ y_n &= \mathbf{R}x_n \end{aligned} \right\} \quad \mathbf{R} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix}$$

- Let  $\mathbf{r} = [r_1, r_2, r_3, \dots, r_9]^T$ , we have

$$y_i = \mathbf{R}x_i \quad \Rightarrow \quad \underbrace{\begin{bmatrix} x_i^T & 0 & 0 \\ 0 & x_i^T & 0 \\ 0 & 0 & x_i^T \end{bmatrix}}_{3 \times 9} \mathbf{r} = y_i$$

# Rotation-only 3D-3D registration



- We can solve the vector of rotation elements by using at least three point correspondences.

$$\begin{bmatrix} \mathbf{x}_1^T & 0 & 0 \\ 0 & \mathbf{x}_1^T & 0 \\ 0 & 0 & \mathbf{x}_1^T \\ \dots & \dots & \dots \\ \mathbf{x}_i^T & 0 & 0 \\ 0 & \mathbf{x}_i^T & 0 \\ 0 & 0 & \mathbf{x}_i^T \\ \dots & \dots & \dots \end{bmatrix} \mathbf{r} = \begin{bmatrix} \mathbf{y}_1 \\ \dots \\ \mathbf{y}_i \\ \dots \end{bmatrix}$$

$$\mathbf{X}\mathbf{r} = \mathbf{y}$$

$$\mathbf{r} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

# Rotation-only 3D-3D registration



- However, DLT method does not impose the  $SO(3)$  constraints on the rotation matrix.

$$\det(\mathbf{R}) = 1, \mathbf{R}^T \mathbf{R} = \mathbf{I}$$

- We can get a rotation that is the closest to the solved  $\mathbf{R}^*$  by SVD:

$$\mathbf{R}^* = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$$

$$\mathbf{R} \leftarrow \mathbf{U} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{V}^T$$

# Rotation-only 3D-3D registration



- **Second approach – Absolute orientation** : using unit quaternion (naturally handling the nonlinear constraints on rotation)
- A vector  $x_i$  after a rotation  $q$  is given by  $y_i$

$$\tilde{y}_i = q \otimes \tilde{x}_i \otimes q^*$$

- Here,  $\tilde{x}_i = \begin{bmatrix} 0 \\ x_i \end{bmatrix} \in \mathbb{R}^{4 \times 1}$  and  $\tilde{y}_i = \begin{bmatrix} 0 \\ y_i \end{bmatrix} \in \mathbb{R}^{4 \times 1}$ . We want to seek the unit quaternion :

$$q \in \mathbb{R}^{4 \times 1}, \|q\| = 1$$



# Rotation-only 3D-3D registration



- All we want to do is to seek a unit quaternion that makes

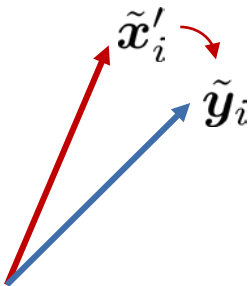
$$\tilde{\mathbf{y}}_i \leftrightarrow \tilde{\mathbf{x}}'_i = \mathbf{q} \otimes \tilde{\mathbf{x}}_i \otimes \mathbf{q}^*$$

- as close as possible.
- A straightforward way is to minimize the sum of squares :

$$\|\tilde{\mathbf{y}}_i - \tilde{\mathbf{x}}'_i\|^2$$

$$\rightarrow \|\tilde{\mathbf{y}}_i\|^2 + \|\tilde{\mathbf{x}}'_i\|^2 - 2(\tilde{\mathbf{y}}_i, \tilde{\mathbf{x}}'_i)$$

*const.*



# Rotation-only 3D-3D registration



- So what we need to do is to find a unit quaternion that maximize the dot products of two quaternions:

$$\begin{aligned} & (\tilde{\mathbf{y}}_i, \tilde{\mathbf{x}}'_i) \\ &= (\tilde{\mathbf{y}}_i, \mathbf{q} \otimes \tilde{\mathbf{x}}_i \otimes \mathbf{q}^*) \end{aligned}$$

- Recall that the dot product of two quaternions is invariant to quaternion multiplication.

$$(\mathbf{q} \otimes \mathbf{r})^T (\mathbf{q} \otimes \mathbf{t}) = ([\mathbf{q}]_L \mathbf{r})^T ([\mathbf{q}]_L \mathbf{t}) = \mathbf{r}^T [\mathbf{q}]_L^T [\mathbf{q}]_L \mathbf{t} = \mathbf{r}^T \mathbf{t}$$

# Rotation-only 3D-3D registration



- We have  $(\tilde{\mathbf{y}}_i, \mathbf{q} \otimes \tilde{\mathbf{x}}_i \otimes \mathbf{q}^*)$   
 $= (\tilde{\mathbf{y}}_i \otimes \mathbf{q}, \mathbf{q} \otimes \tilde{\mathbf{x}}_i)$   
 $= ([\tilde{\mathbf{y}}_i]_L \mathbf{q})^T ([\tilde{\mathbf{x}}_i]_R \mathbf{q})$   
 $= \mathbf{q}^T [\tilde{\mathbf{y}}_i]_L^T [\tilde{\mathbf{x}}_i]_R \mathbf{q}$

- We want to solve the maximization problem

$$\arg \max_{\mathbf{q}} \mathbf{q}^T \mathbf{A} \mathbf{q}$$

with respect to  $\mathbf{q}$ ,  $\|\mathbf{q}\| = 1$ , here  $\mathbf{A} = [\tilde{\mathbf{y}}_i]_L^T [\tilde{\mathbf{x}}_i]_R$

# Rotation-only 3D-3D registration



- Let  $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4$ ,  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$  be the eigenvectors and corresponding eigenvalues of  $\mathbf{A}$ .

$$\lambda_i \mathbf{q}_i = \mathbf{A} \mathbf{q}_i \quad (\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \lambda_4)$$

- The eigenvectors of a symmetric matrix are orthogonal:

$$\mathbf{q}_i^T \mathbf{q}_j = 0, (i \neq j), \|\mathbf{q}_i\| = 1$$

- $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4\}$  spans the eigen space.



# Rotation-only 3D-3D registration



- For a given unit quaternion, we can represent it within the eigen space

$$\mathbf{q} = a_1 \mathbf{q}_1 + a_2 \mathbf{q}_2 + a_3 \mathbf{q}_3 + a_4 \mathbf{q}_4$$


- where  $a_1^2 + a_2^2 + a_3^2 + a_4^2 = 1$ .

  $(\lambda_i \mathbf{q}_i = \mathbf{A} \mathbf{q}_i)$

$$\mathbf{A} \mathbf{q} = a_1 \lambda_1 \mathbf{q}_1 + a_2 \lambda_2 \mathbf{q}_2 + a_3 \lambda_3 \mathbf{q}_3 + a_4 \lambda_4 \mathbf{q}_4$$



$$\mathbf{q}^T \mathbf{A} \mathbf{q} = a_1^2 \lambda_1 + a_2^2 \lambda_2 + a_3^2 \lambda_3 + a_4^2 \lambda_4$$

  $(a_1^2 = 1 - a_2^2 - a_3^2 - a_4^2)$

$$\mathbf{q}^T \mathbf{A} \mathbf{q} = \lambda_1 + a_2^2(\lambda_2 - \lambda_1) + a_3^2(\lambda_3 - \lambda_1) + a_4^2(\lambda_4 - \lambda_1)$$



$$\max(\mathbf{q}^T \mathbf{A} \mathbf{q}) = \lambda_4, \mathbf{q} = \mathbf{q}_4$$

# Summary



- Rotation-only 3D-3D registration

- Iterative approach (Gauss-Newton)

$$\Delta\theta = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{z}, \quad \mathbf{R} \leftarrow \mathbf{R} \exp(\Delta\theta^\wedge)$$

- Close form approach

- Rotation matrix – Direct Linear Transformation

$$\mathbf{X} \mathbf{r} = \mathbf{y} \rightarrow \mathbf{R} \quad (\text{SVD approximation})$$

- Unit quaternion

$$\arg \max_{\mathbf{q}} \mathbf{q}^T \mathbf{A} \mathbf{q}$$

- Usually we first run the close form approach and then refine the estimation by iterative approach

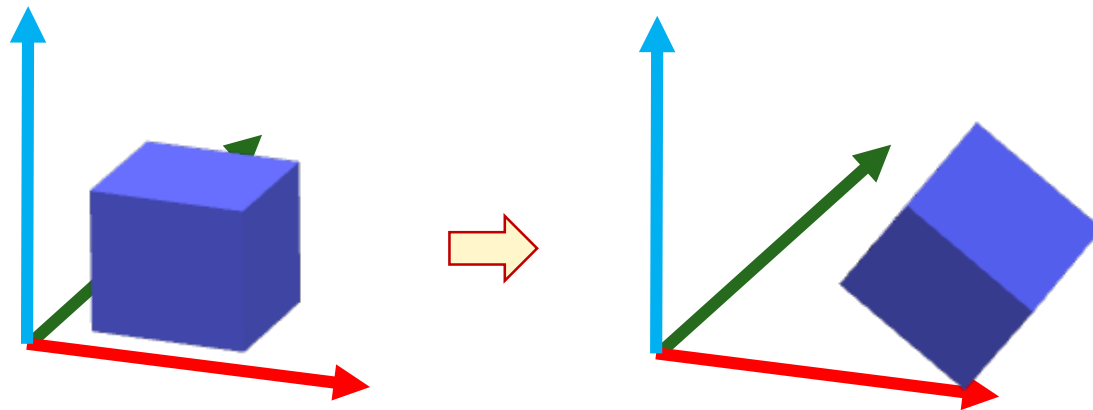
# 3D-3D registration (Rigid)



- Now we consider the 3D-3D registration problem using rigid transformation.

$$\mathbf{y}_i = \mathbf{R}\mathbf{x}_i + \mathbf{t}$$

$(i = 1, 2, \dots)$



## 3D-3D registration (Rigid)



- We can manage to obtain  $\mathbf{R}$  by rotation-only registration

$$\mathbf{y}_1 = \mathbf{R}\mathbf{x}_1 + \mathbf{t}$$

$$\mathbf{y}_2 = \mathbf{R}\mathbf{x}_2 + \mathbf{t} \quad \Rightarrow \quad \bar{\mathbf{y}} = \mathbf{R}\bar{\mathbf{x}} + \mathbf{t}$$

... ..

$$\mathbf{y}_n = \mathbf{R}\mathbf{x}_n + \mathbf{t}$$



$$(\mathbf{y}_1 - \bar{\mathbf{y}}) = \mathbf{R}(\mathbf{x}_1 - \bar{\mathbf{x}})$$

$$(\mathbf{y}_2 - \bar{\mathbf{y}}) = \mathbf{R}(\mathbf{x}_2 - \bar{\mathbf{x}}) \quad \Rightarrow \quad \mathbf{r}_i = \mathbf{R}\mathbf{s}_i$$

... ..

$$(\mathbf{y}_n - \bar{\mathbf{y}}) = \mathbf{R}(\mathbf{x}_n - \bar{\mathbf{x}})$$



# 3D-3D registration (Rigid)



- **3D-3D registration algorithm**
- Inputs : corresponding 3D points  $\mathbf{x}_i \leftrightarrow \mathbf{y}_i, (i = 1, \dots, n)$
- Outputs : rigid transformation  $\mathbf{R}, t$ 
  - Step 1. Compute the normalized vectors

$$\mathbf{s}_i = \mathbf{x}_i - \bar{\mathbf{x}}, \mathbf{r}_i = \mathbf{y}_i - \bar{\mathbf{y}}$$

- Step 2. Get the orientation by DLT or absolute orientation algorithm

$$\arg \max_{\mathbf{q}} \mathbf{q}^T \mathbf{A} \mathbf{q} \rightarrow \mathbf{q} \text{ or } \mathbf{X} \mathbf{r} = \mathbf{y} \rightarrow \mathbf{R}$$

- Step 3. Compute the translation vector

$$t = \bar{\mathbf{y}} - \mathbf{R} \bar{\mathbf{x}}$$

## 3D-3D registration (Rigid)



- We can minimize the nonlinear least squares to refine the estimation.

$$\arg \max_{\mathbf{R}, \mathbf{t}} \sum_{i=1}^n \|\mathbf{y}_i - \mathbf{R}\mathbf{x}_i - \mathbf{t}\|^2$$

- Gauss-Newton or Levenberg-Marquardt algorithm can be applied.

$$\mathbf{X} \leftarrow \mathbf{X} \boxplus \Delta \mathbf{x}$$

$$\begin{pmatrix} \mathbf{R} \\ \mathbf{t} \end{pmatrix} \leftarrow \begin{pmatrix} \mathbf{R} \exp(\Delta \theta^\wedge) \\ \mathbf{t} + \Delta \mathbf{t} \end{pmatrix}$$

- All we need is to solve the incremental step  $\Delta \mathbf{x} = \begin{pmatrix} \Delta \theta \\ \Delta \mathbf{t} \end{pmatrix}$

# 3D-3D registration (Rigid)



- Using the first-order approximation :

$$\begin{aligned} & \sum_{i=1}^n \| \mathbf{y}_i - \mathbf{R} \exp(\Delta \theta^\wedge) \mathbf{x}_i - \mathbf{t} - \Delta \mathbf{t} \|^2 \\ & \approx \sum_{i=1}^n \| \mathbf{y}_i - \mathbf{R}(\mathbf{I} + [\Delta \theta]_\times) \mathbf{x}_i - \mathbf{t} - \Delta \mathbf{t} \|^2 \\ & = \sum_{i=1}^n \| \mathbf{y}_i - \mathbf{R} - \mathbf{t} + [\mathbf{R} \mathbf{x}_i]_\times \Delta \theta - \Delta \mathbf{t} \|^2 \\ & \quad \sum_{i=1}^n \| \mathbf{z}_i - \mathbf{J}_i \begin{bmatrix} \Delta \theta \\ \Delta \mathbf{t} \end{bmatrix} \|^2 \end{aligned}$$

# 3D-3D registration (Rigid)



- Gauss-Newton :

$$\Delta \mathbf{x} = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{z}$$

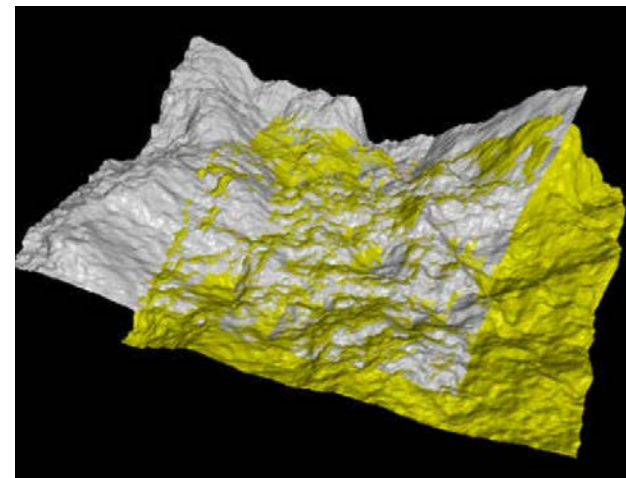
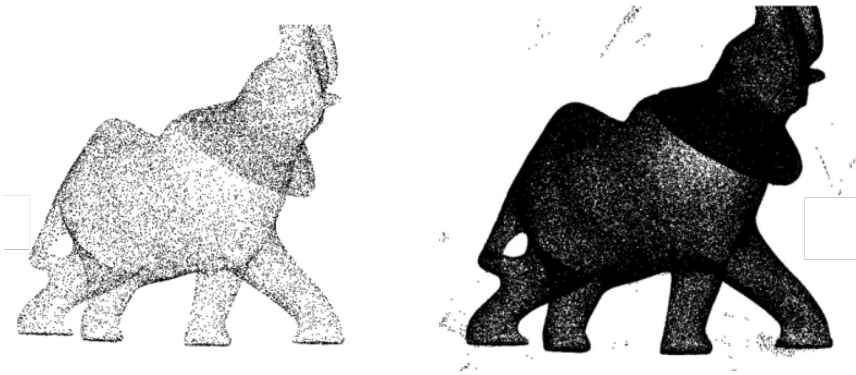
- Levinberg-Marquardt :

$$\Delta \mathbf{x} = (\mathbf{J}^T \mathbf{J} + \lambda \mathbf{I})^{-1} \mathbf{J}^T \mathbf{z}$$

# Iterative Closest Point algorithm



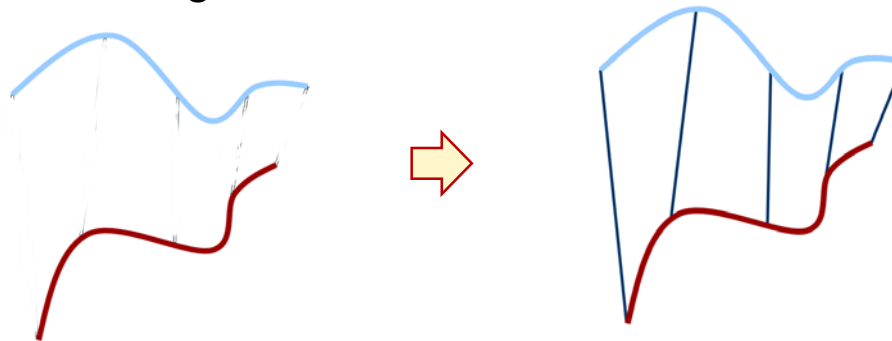
- If we do not know the point correspondences, how do we get the rigid transformation ?
- For example, given two 3D LiDAR scans
  - The density is different
  - Only the parts of the scans are overlapped



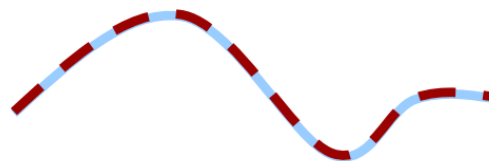
# Iterative Closest Point algorithm



- Key idea: Iterate to find the corresponding points and estimate the rigid transformation (alignment)
  - Step 1 – **Matching** : find the closest point as the corresponding point using the current alignment.



- Step2 – **Updating** : compute the alignment using the close-form solution as introduced previously.



# Iterative Closest Point algorithm



- Matching step : for each point  $y_i$  in the point cloud, we search its corresponding point  $x_i^*$  using the current transformation

$$\arg \min_j \|y_i - x'_j\|^2 \rightarrow x_i^*$$
$$(x'_j = \mathbf{R}x_j + t)$$

- After that, we get a set of corresponding 3D points

$$y_1 \leftrightarrow x_1^*$$

$$y_2 \leftrightarrow x_2^*$$

...

$$y_n \leftrightarrow x_n^*$$



$\mathbf{R}, t$

1. DLT (matrix)

2. Absolute orientation (quaternion)

3. Gauss-Newton/Levenberg-Marquardt



# Iterative Closest Point algorithm



- ICP converges only if the starting point is “close enough” to the real solution.



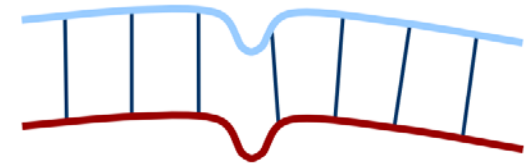
- The performance (accuracy & efficiency ) largely depends on the first step – **matching**.
- In practice, several techniques can be used to improve the **matching** performance.

# Iterative Closest Point algorithm

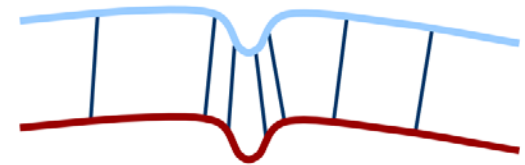


## ▪ Sampling

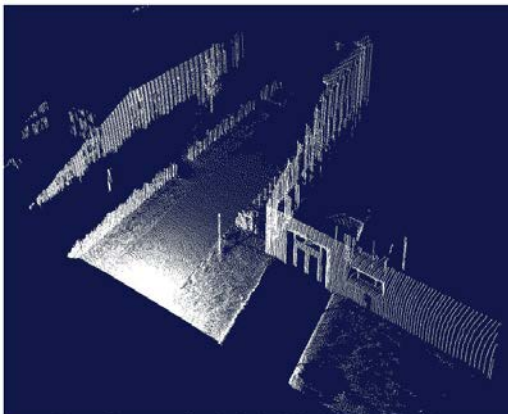
- Uniform sampling
- Random sampling
- Normal-space sampling (use when the normal is available)
- Feature-based sampling



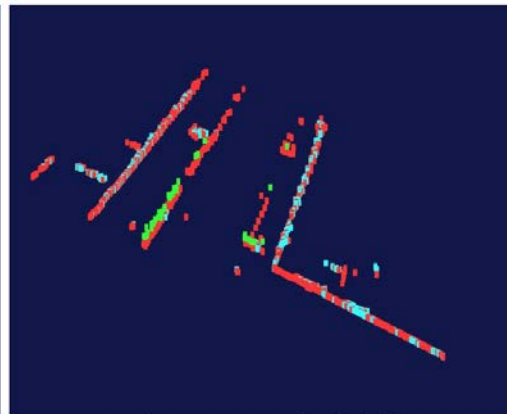
uniform sampling



normal-space sampling



3D Scan (~200.000 Points)

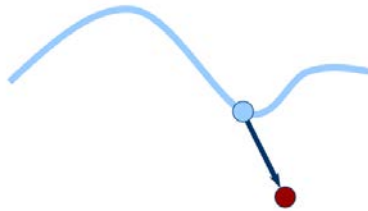


Extracted Features (~5.000 Points)

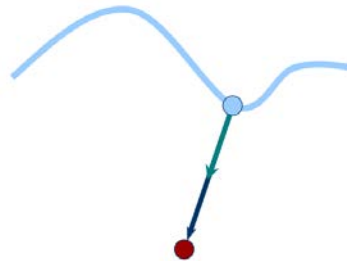
# Iterative Closest Point algorithm



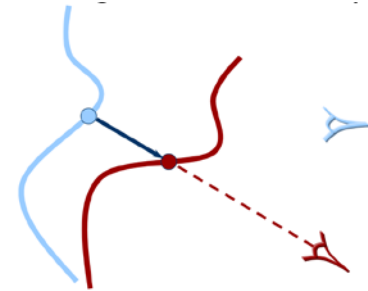
- Finding the corresponding point



Closest point



Normal shooting



Projection

- Using KD-trees or Oct-trees to improve the matching speed.

# Summary



- **ICP** algorithm aligns two points clouds without known their correspondences.
- The two point clouds can be of different density, partially overlapped.
- The key idea is to iteratively finding the correspondences and update the rigid transformation.
- A good matching strategy is critical to the performance.
- A initial guess is also critical.
- **ICP** algorithm is widely used in autonomous cars who equipped with LiDAR.

# Outline



- ~~About pose estimation~~
- ~~3D-to-3D registration~~
  - ~~Rotation only~~
  - ~~Rotation plus Translation~~
  - ~~Unknown correspondences – Iterative Closest Point (ICP)~~
- **3D-to-2D registration (Camera pose estimation)**
  - 3D objects
    - Close-form algorithm - P3P
    - Iterative algorithm - POSIT
    - Pose refinement - Nonlinear least squares
  - Planar objects
    - Known patterns – Checkboard box, QR pattern
    - Planar Pictures