FCV2010 P2-20 Hiroshima University

Comparison of 3DOF Pose Representations For Pose Estimation

Kengo Harada† Satoko Tanaka† Toru Tamaki† Bisser Raytchev† Kazufumi Kaneda† Toshiyuki Amano‡
†: Hiroshima University, Japan ‡: NAIST, Japan

Linear Pose Estimation

Training

Training Images  $\boldsymbol{x}_1$ $\boldsymbol{x}_2$ $\boldsymbol{x}_3$ $\boldsymbol{x}_4$ …

Pose Parameters $\boldsymbol{p}_1$ $\boldsymbol{p}_2$ $\boldsymbol{p}_3$ $\boldsymbol{p}_4$ …

$$\boldsymbol{p}_j = F\boldsymbol{x}_j$$

Estimation

 $\boldsymbol{x}$? ? → $\boldsymbol{p} = F\boldsymbol{x}$

Question

Is estimation error of rotation matrix smaller than other representations ??

Pose Estimation as Approximate and Two Properties

An image can be approximated by their linear combination of training images $x \simeq \sum b_j \boldsymbol{x}_j$

Example of a 1DOF case

A pose estimate is represented by a linear combination of training poses $\boldsymbol{p} \simeq F\boldsymbol{x} = \sum b_j F\boldsymbol{x}_j = b_j \boldsymbol{p}_j$

A pose should be bijective with appearance

$\boldsymbol{p}_1 \neq \boldsymbol{p}_2$

$\boldsymbol{p}_1 = F\boldsymbol{x}$

$\boldsymbol{p}_2 = F\boldsymbol{x}$

F **does not exist**

A pose should be continuous with appearance

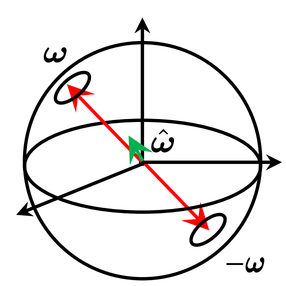
b_{k+1} and b_{k+2} have larger values than others if training is successful

Parameters		
θ	$180° \simeq \cdots + b_k \cdot 340° + b_{k+1} \cdot 350° + b_{k+2} \cdot 10° + b_{k+3} \cdot 20° + \cdots$	×
$\begin{bmatrix}\sin\theta \\ \cos\theta\end{bmatrix}$	$\begin{cases}\sin 0° \simeq \cdots + b_k \sin 340° + b_{k+1} \sin 350° + b_{k+2} \sin 10° + b_{k+3} \sin 20° + \cdots \\ \cos 0° \simeq \cdots + b_k \cos 340° + b_{k+1} \cos 350° + b_{k+2} \cos 10° + b_{k+3} \cos 20° + \cdots\end{cases}$	○

Pose Representations and Properties

Representation	Parameters	Bijection	Continuity
Rotation matrix	$[r_{11}\ r_{12}\ \cdots\ r_{33}]^T$	○	○
ZYX Euler angles	$[\theta_x\ \theta_y\ \theta_z]^T$ $(-\pi \leq \theta_{x,y,z} < \pi)$	×	×
Exponential map	$\boldsymbol{\omega} = [\omega_1\ \omega_2\ \omega_3]^T$ $(0 \leq \lvert\boldsymbol{\omega}\rvert \leq \pi)$	×	×
Unit quaternions	$\boldsymbol{q} = [q_0\ q_1\ q_2\ q_3]^T$	×	○

Exponential map

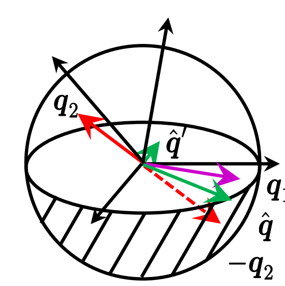


Sphere in 3-dimentional space $r = \pi$ [rad]

- $\boldsymbol{\omega}_j$ is on a sphere with radius π [rad]
- When $\|\boldsymbol{\omega}\| = \pi$ [rad], $\boldsymbol{\omega}$ and $-\boldsymbol{\omega}$ are the same pose

Not bijective

Unit quaternions



Unit hyper-sphere in 4-dimentional space

- $\boldsymbol{q}_j$ is on a unit sphere in 4-dimentional space
- $\boldsymbol{q}$ and $-\boldsymbol{q}$ are the same pose (non bijective)
- If we discard $-\boldsymbol{q}$, it is bijective

Discontinuity at the edge of the hyper-hemisphere

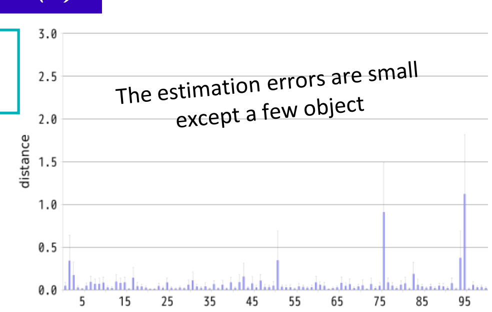
Estimation Method and Error Metric

1. Given training images $\boldsymbol{x}_j$ and pose parameters $\boldsymbol{p}_j$ in a vector, the equations are stacked:
$$\boldsymbol{p}_j = F\boldsymbol{x}_j$$
2. The pose of a test image $\boldsymbol{x}$ is estimated by:
$$\boldsymbol{p} = F\boldsymbol{x}$$
3. Normalization:
   - For rotation matrix, $\boldsymbol{p} \rightarrow 3 \times 3$ matrix, and it is converted to a rotation matrix by using polar decomposition.
   - For unit quaternions, $\|\boldsymbol{p}\| = 1$
4. Conversion to a rotation matrix: Estimated poses are converted to corresponding rotation matrix: $\hat{R}$
5. Error as distance between rotation matrices: R_t is a true rotation matrix

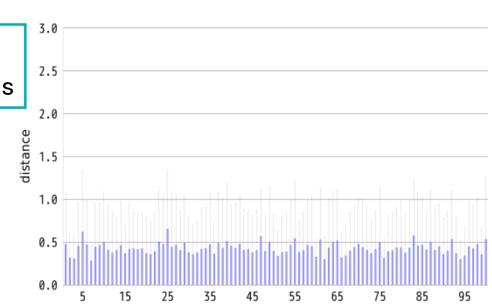
$$d_F\left(R_t, \hat{R}\right) = \frac{1}{\sqrt{2}}\left\|\log R_t \hat{R}\right\|, \quad \log R = \begin{cases} 0, & (\theta = 0) \\ \frac{\theta}{2\sin\theta}\left(R - R^t\right), & (\theta \neq 0)\end{cases}$$

Results (1)

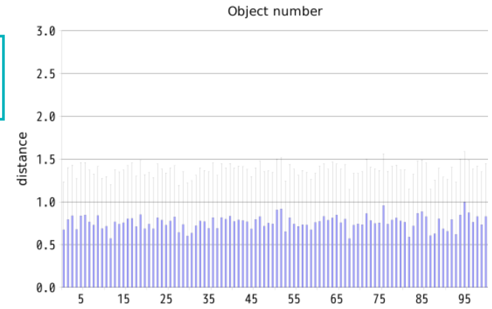
Rotation matrix



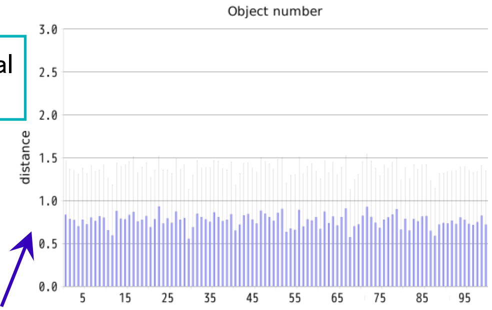
Unit quaternions



Euler angles



Exponential map



The errors are big averagely

Experimental Setup

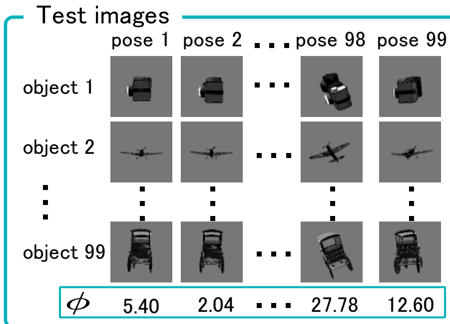
We focused poses around discontinuity: some pose representations have discontinuity at a rotation angle of π.

Images are created as follows:

- Create a rotation matrix R_z with a rotation about z axis by π
- Create a small random rotation R_s with a rotation about a random axis by an angle ϕ uniformly distributed in $[0, \pi/6]$
- Combine them: a (true) rotation matrix is
$$R_t = R_s R_z$$

- We use 100 3D objects
- 2500 training images
- 100 test images

Test images

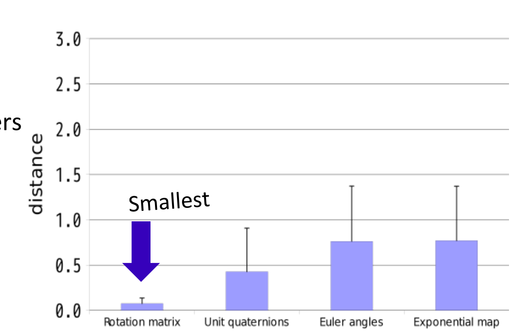


Results (2)

- Pairwise *t*-test between rotation matrix and the others

$(p < 0.01)$

All tests are significantly different.



The average of the distance