On the Continuity of Rotation Representations in Neural Networks

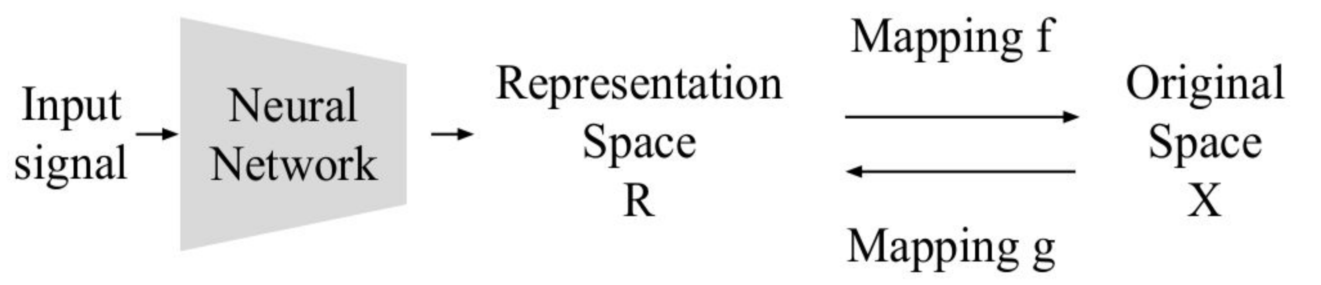
Yi Zhou*[1] Connelly Barnes*[2] Jingwan Lu[2] Jimei Yang[2] Hao Li[1, 3, 4]

[1]University of Southern California [2]Adobe Research [3]USC Institute for Creative Technologies [4]Pinscreen

Why does my network still occasionally produce big errors even with abundant weights and extensive training?

Continuous representations in neural networks:

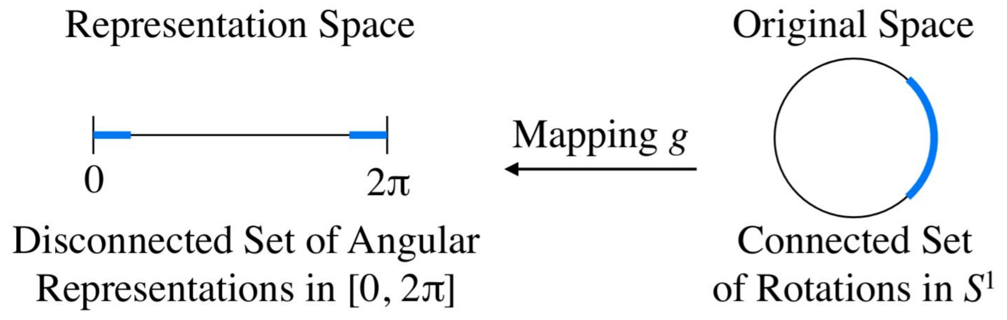
- The mapping g in the below diagram needs to be continuous.
- In other words, g is an embedding (X is homeomorphic to a subset of R)



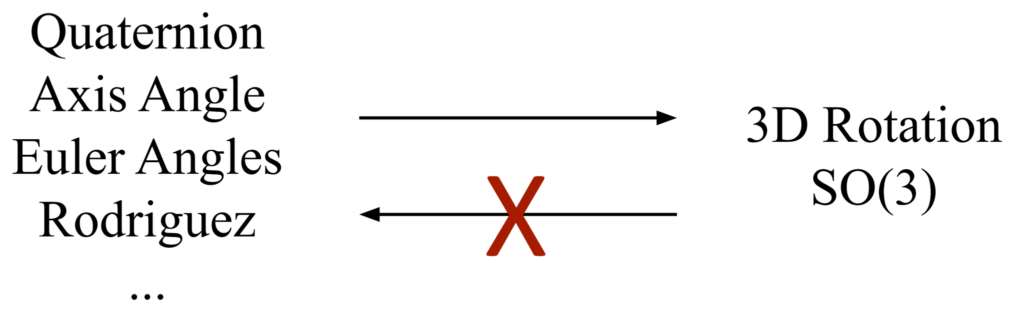
- Discontinuous representations are hard for neural networks to approximate.

Commonly used rotation representations are **discontinuous**.

- A 2D example: Polar angle representation



Note: For 3D rotations, all representations are discontinuous in the real Euclidean spaces of **four or fewer** dimensions.



- Be careful when using 4D or fewer dimensional rotation representations because they have discontinuities which manifest as degenerations (e.g. Gimbal lock or sign flips) when dealing with the whole 3D rotation space. If you must use a discontinuous representation then try to parameterize it so your training set avoids the discontinuities.

Let's find **continuous** 3D rotation representations.

6D representation for 3D rotations

Mapping from SO(3) (matrix) to a 6D representation (Stiefel Manifold):

$$g_{GS}\left(\begin{bmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{bmatrix}\right) = \begin{bmatrix} | & | \\ a_1 & a_2 \\ | & | \end{bmatrix}$$

Mapping from 6D representation to SO(3) (Gram-Schmidt-like orthogonalization):

$$f_{GS}\left(\begin{bmatrix} | & | \\ a_1 & a_2 \\ | & | \end{bmatrix}\right) = \begin{bmatrix} | & | & | \\ b_1 & b_2 & b_3 \\ | & | & | \end{bmatrix}$$

$$b_i = \left[\begin{cases} N(a_1) & \text{if } i = 1 \\ N(a_2 - (b_1 \cdot a_2) b_1) & \text{if } i = 2 \\ b_1 \times b_2 & \text{if } i = 3 \end{cases}\right]^T$$

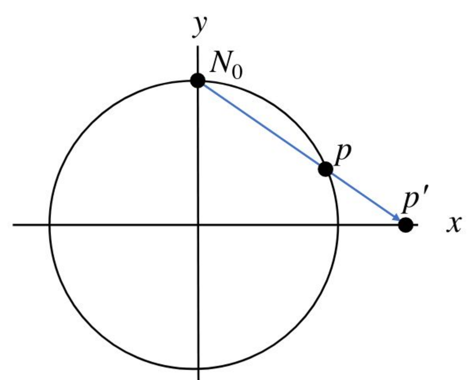
$N(v) = v \,/\, \| v \|$ (Normalization)

5D representation for 3D rotations

Eliminating one dimension of g_{GS} by stereographically projecting $[a_{13}, a_{21}, a_{22}, a_{23}]^T$ to 3D.

6D: a_{11} a_{12} a_{13} a_{21} a_{22} a_{23}

5D: a_{11} a_{12} c_1 c_2 c_3

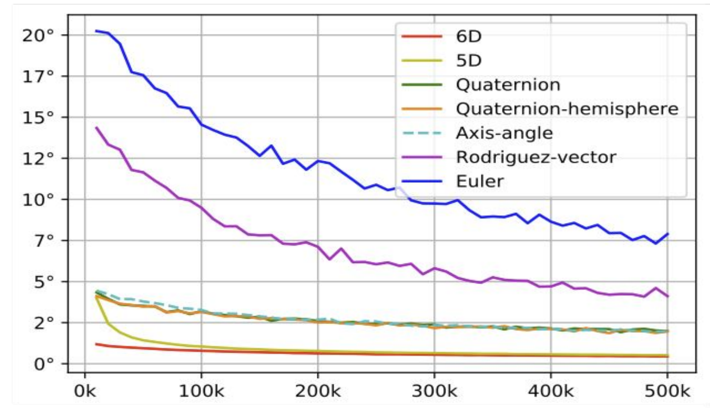


An illustration of stereographic projection in 2D. We are given as input a point p on the unit sphere S^1. We construct a ray from a fixed projection point $N_0 = (0, 1)$ through p and find the intersection of this ray with the plane y = 0. The resulting point p' is the stereographic projection of p.

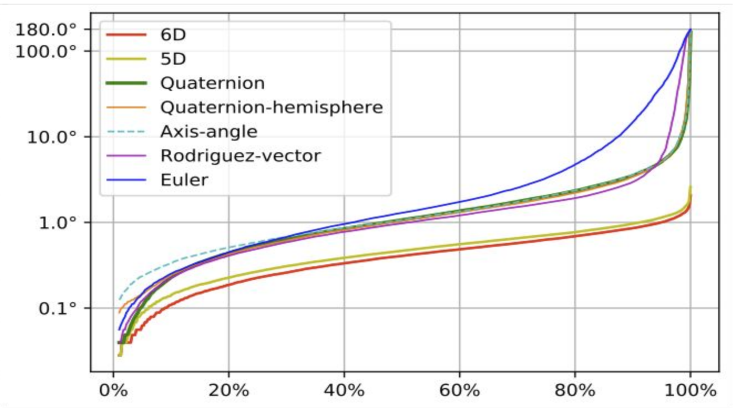
- For the general case of the n dimensional rotation group SO(n), please check our paper. We also discuss other groups such as the orthogonal group and similarity transforms.

Sanity Test:

Sanity Test – Geodesic Loss and Uniform Sampling on Axis and Angle



a. Mean errors during iterations.

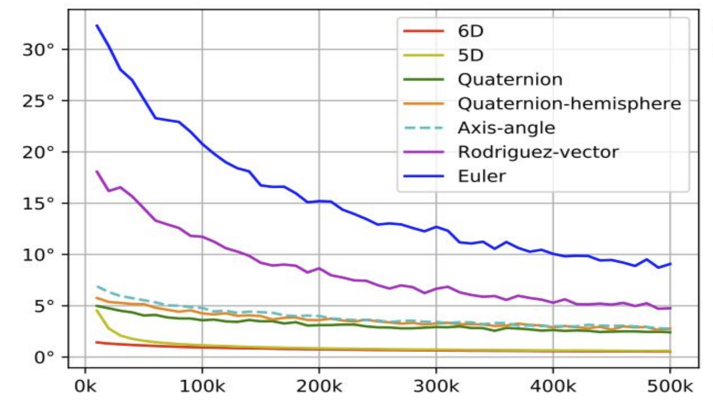


b. Percentile of errors at 500k iteration.

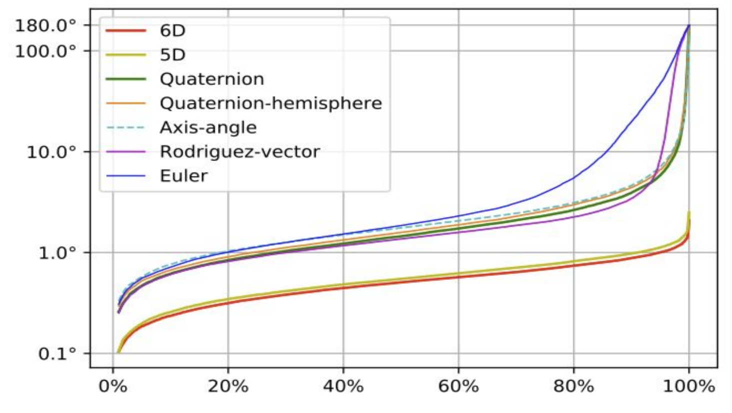
	Mean(°)	Max(°)	Std(°)
6D	**0.45**	**2.33**	**0.29**
5D	0.52	3.19	0.33
Quat	2.02	180.0	6.08
Quat-hemi	3.47	179.22	6.72
AxisA	2.0	179.58	5.43
Euler	7.45	179.87	21.87
Rodriguez	4.17	180.0	17.78

c. Errors at 500k iteration.

Sanity Test – Geodesic Loss and Uniform Sampling on SO(3)



d. Mean errors during iterations.



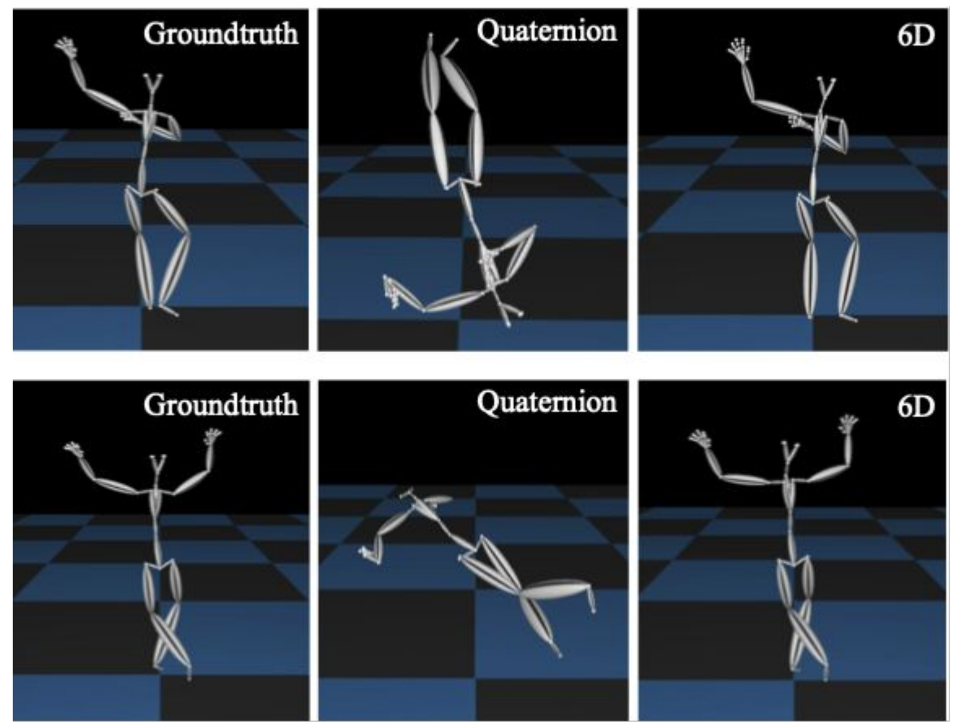
e. Percentile of errors at 500k iteration.

	Mean(°)	Max(°)	Std(°)
6D	**0.54**	**1.82**	**0.26**
5D	0.59	1.93	0.29
Quat	2.4	176.9	6.81
Quat-hemi	2.72	179.55	7.03
AxisA	2.89	175.53	6.77
Euler	9.12	179.93	24.01
Rodriguez	4.74	179.45	17.84

f. Errors at 500k iteration.

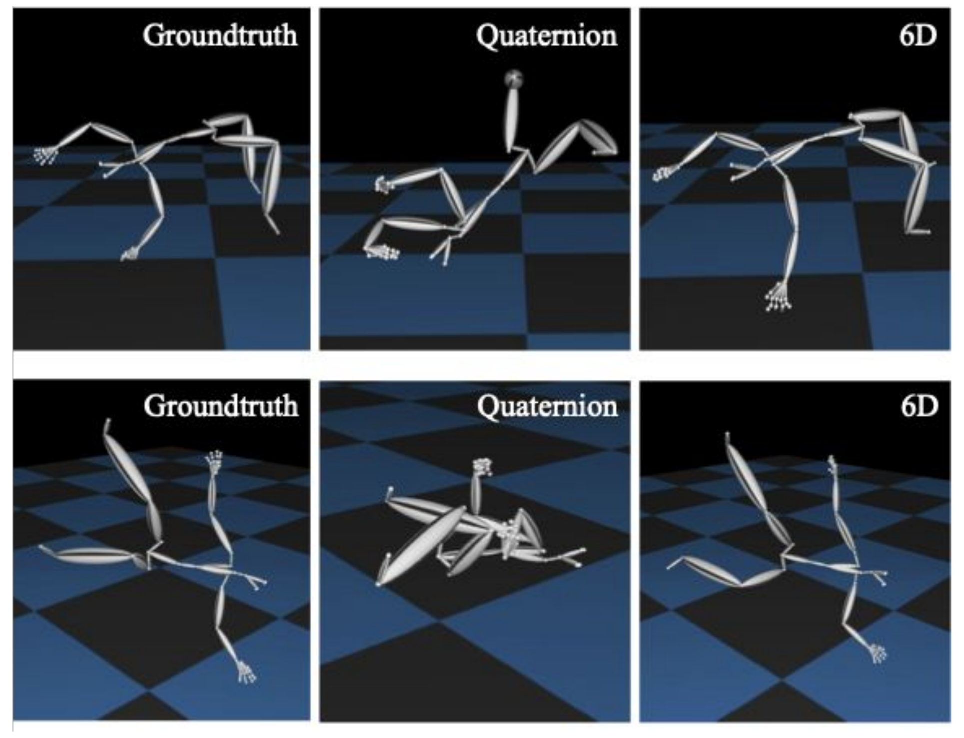
Inverse Kinematics Test:

Worst two frames using quaternions.



IK results for the two frames with highest pose error from the test set for the network trained using quaternions, and the corresponding results on the same frames for the network trained on the 6D representation.

Worst two frames using 6D representations.



IK results for the two worst frames for the 6D representation network, and the corresponding results for the quaternion network.