PDO-s3DCNNs: Partial Differential Operator Based Steerable 3D CNNs

Zhengyang Shen, Tao Hong, Qi She, Jinwen Ma, Zhouchen Lin

Peking University

Introduction

- An extension work of our previous two works: PDO-eConvs[1] and PDO-eS2CNNs [2].
- As far as we know, PDO-s3DCNNs are the most general and flexible steerable 3D CNNs.

[1] Shen et al. PDO-eConvs: Partial Differential Operator Based Equivariant Convolutions. ICML, 2020.

[2] Shen et al. PDO-eS2CNNs: Partial Differential Operator Based Equivariant Spherical CNNs. AAAI, 2021.

Steerable CNNs

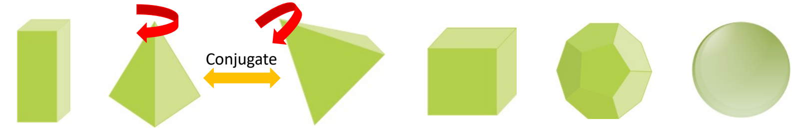
For a 3D steerable CNN Ψ, it should satisfy that:

$$\forall g \in \mathcal{G}, \quad \pi'(g)\left[\Psi\left[f\right]\right] = \Psi\left[\pi(g)[f]\right],$$

where

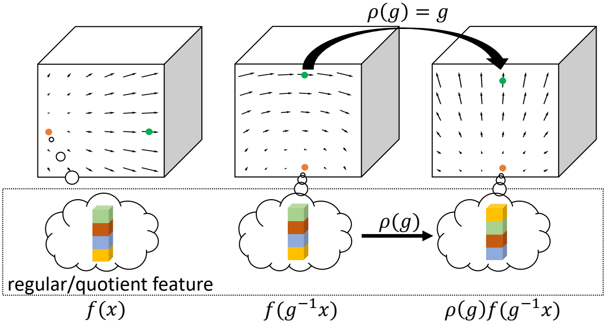
$$\left[\pi(g)f\right](x) = \rho(g)f\left(g^{-1}x\right),$$

As a result, the most general steerable CNNs should address arbitrary rotation group $\mathcal{G}$ and their feature field determined by $\rho(g)$.



(a) Rectangle (b) Tetrahedron (c) Cube (d) Icosahedron (e) Sphere

Common 3D rotation groups contain the continuous group SO(3) and its discrete subgroups, including the dihedral group D_N, the tetrahedral group $\mathcal{T}$, the cubic group $\mathcal{C}$ and the icosahedral group $\mathcal{I}$.



Common feature fields include regular features, quotient features and irreducible features.

Table: The comparison between PDO-s3DCNNs and other 3D steerable models.

	$\mathcal{G}$			Feature field			
	$\mathcal{G} \leq \mathcal{O}$	$\mathcal{I}$	$SO(3)$	Regular	Quotient	Irreducible	Data type
N-Body			✓			✓	graphs
TFN			✓			✓	point clouds
CubeNets	✓			✓			voxels
SE3CNNs			✓			✓	voxels
SE(3)-Transformer			✓			✓	point clouds/graphs
PDO-s3DCNN	✓	✓	✓	✓	✓	✓	voxels

PDO-s3DCNNs can accommodate all common subgroups of $SO(3)$ and feature fields, while others can only address specific groups and feature fields.

Main theoretical results

We employ a combination of PDOs to define a 3D filter on the input function $f \in C^\infty(\mathbb{R}^3, \mathbb{R}^K)$:

$$\Psi[f] = \Big(A_0 + A_1\partial_{x_1} + A_2\partial_{x_2} + A_3\partial_{x_3} + A_{11}\partial_{x_1^2} + A_{12}\partial_{x_1 x_2} + A_{13}\partial_{x_1 x_3} + A_{22}\partial_{x_2^2} + A_{23}\partial_{x_2 x_3} + A_{33}\partial_{x_3^2} \Big)[f].$$

Ψ is equivariant over $\mathcal{G}$, if and only if its coefficients satisfy the following linear constraints: $\forall g \in \mathcal{G}$,

$$\begin{cases} r\rho'(g)B_0 = B_0\rho(g), \\ \rho'(g)B_1 = B_1\left(g \otimes \rho(g)\right), \\ \rho'(g)B_2 = B_2\left(P\left(g \otimes g\right)P^\dagger \otimes \rho(g)\right), \end{cases}$$

$$
\begin{cases}
\left(I_K \otimes \rho'(g) - \rho(g)^T \otimes I_{K'}\right) vec(B_0) = 0, \\
\left(I_{3K} \otimes \rho'(g) - (g \otimes \rho(g))^T \otimes I_{K'}\right) vec(B_1) = 0, \\
\left(I_{6K} \otimes \rho'(g) - \left(P(g \otimes g)P^{\dagger} \otimes \rho(g)\right)^T \otimes I_{K'}\right) vec(B_2) = 0,
\end{cases}
$$

$$
\begin{cases}
B_0 = A_0, \\
B_1 = \left[A_1, A_2, A_3\right], \\
B_2 = \left[A_{11}, A_{12}, A_{13}, A_{22}, A_{23}, A_{33}\right],
\end{cases}
$$

$$
P = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix},
$$

- The linear constraints can be solved efficiently using SVD.
- For discrete groups, we should only solve the constraints for group generators.
- For continuous group $SO(3)$, we should only solve the constraints for approximate group generators.

Experimental results

Rotated Tetris:

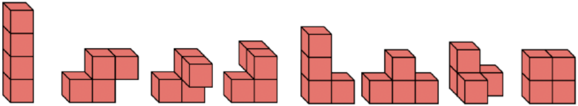


Table: The test accuracy of the $\mathcal{O}$- and $SO(3)$- steerable CNNs discretized by FD on 3D Tetris with cubic rotations.

Group	Feature field	Test acc. (%)	# Params	Time
$\mathcal{O}$	Regular	100.0 ± 0.0	31k	14.3s
$\mathcal{O}$	$\mathcal{V}$-quotient	100.0 ± 0.0	5.5k	2.3s
$\mathcal{O}$	$\mathcal{T}$-quotient	100.0 ± 0.0	2.2k	1.3s
$SO(3)$	Irreducible	100.0 ± 0.0	22.8k	66.7s

SHREC'17 Retrieval:

Table: The retrieval performance of $\mathcal{V}$-, $\mathcal{T}$-, $\mathcal{O}$-, $\mathcal{I}$- and $SO(3)$-steerable CNNs, tested on SHREC'17.

Group	Discretization	Feature field	Score
$\mathcal{V}$	FD	Regular	52.7
$\mathcal{T}$	FD	Regular	57.6
$\mathcal{O}$	FD	Regular	**58.6**
$\mathcal{I}$	Gaussian	Regular	55.5
$SO(3)$	FD	Irreducible	57.4
$SO(3)$	Gaussian	Irreducible	58.3

Table: The comparison with other equivariant methods on SHREC' 17.

Methods	Score	micro P@N	micro R@N	micro mAP	macro P@N	macro R@N	macro mAP	Param	Input
RI-GCN	56.2	69.1	68.0	64.5	47.4	57.0	47.8	4.4M	point clouds
Li et al.	56.5	69.4	69.4	65.8	48.1	56.0	47.2	2.9M	point clouds
S2CNN	-	70.1	71.1	67.6	-	-	-	1.4M	spherical
FFS2CNN	-	70.7	72.2	68.3	-	-	-	-	spherical
VolterraNet	-	71.0	70.0	67.0	-	-	-	0.4M	spherical
Esteves et al.	56.5	71.7	**73.7**	68.5	45.0	55.0	44.4	0.5M	spherical
Cobb et al.	-	71.9	71.0	67.9	-	-	-	0.25M	spherical
SE3CNN	55.5	70.4	70.6	66.1	49.0	54.9	44.9	0.14M	spherical
Ours ($SO(3)$)	58.3	73.1	73.4	69.3	**52.5**	55.4	47.3	0.15M	spherical
Ours ($\mathcal{C}$, regular)	58.6	72.9	73.0	68.8	51.9	57.7	48.3	0.15M	spherical
Ours ($\mathcal{C}$, $\mathcal{V}$-quotient)	55.5	69.2	69.6	65.0	48.0	56.3	46.0	0.15M	spherical
Ours ($\mathcal{C}$, mixed)	**59.1**	**73.2**	73.3	**69.3**	51.7	**57.8**	**48.8**	0.15M	spherical

ISBI 2012 segmentation

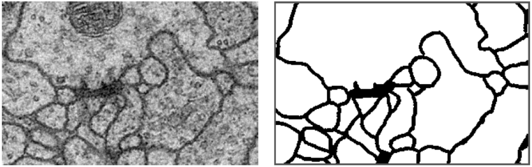


Table: ISBI 2012 segmentation results.

	V_{rand}	V_{info}
U-Net	0.97276	0.98662
FusionNet	0.97804	0.98995
CubeNet	0.98018	0.98202
SFCNN	0.98680	**0.99144**
PDO-s3DCNN ($\mathcal{V}$-steerable)	0.98415	0.99031
PDO-s3DCNN ($\mathcal{C}$-steerable)	**0.98727**	0.99089

Thank you for your time!