Scalable and Equivariant Spherical CNNs by Discrete-Continuous (DISCO) Convolutions

arXiv:2209.13603

Jeremy Ocampo, Matthew A. Price, & Jason D. McEwen*

* Corresponding author: jason.mcewen@kagenova.com, www.jasonmcewen.org

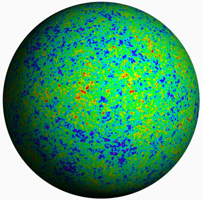
Kagenova Limited, UCL

www.kagenova.com, www.ucl.ac.uk

International Conference on Learning Representations (ICLR) 2023

Data on the sphere

Whenever observe over angles, recover data on 2D sphere (or 3D rotation group).



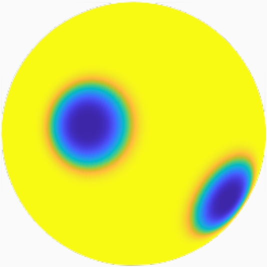
Cosmic microwave background

360° photos/video

Construct CNNs natively on the sphere and encode rotational equivariance.

Categorization of spherical CNNs frameworks

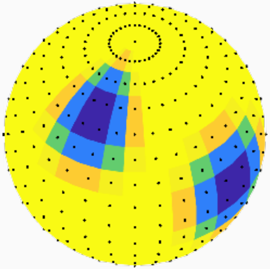
Continuous



✔ Equivariant
✖ Not Scalable

(Cohen et al. 2018, Esteves et al. 2018, Kondor et al. 2018, Cobb et al. 2021, McEwen et al. 2022, ...)

Discrete

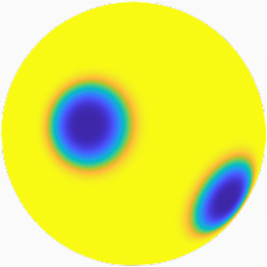


✖ Not Equivariant
✔ Scalable

(Jiang et al. 2019, Zhang et al. 2019, Perraudin et al. 2019, Cohen et al. 2019, ...)

Categorization of spherical CNNs frameworks

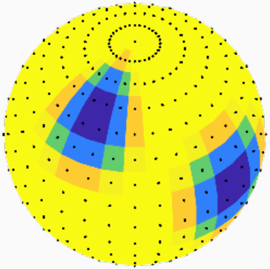
Continuous



✔ Equivariant
✖ Not Scalable

(Cohen et al. 2018, Esteves et al. 2018, Kondor et al. 2018, Cobb et al. 2021, McEwen et al. 2022, ...)

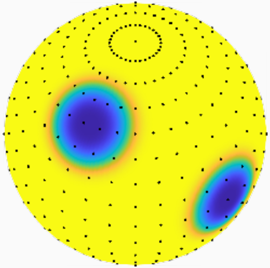
Discrete



✖ Not Equivariant
✔ Scalable

(Jiang et al. 2019, Zhang et al. 2019, Perraudin et al. 2019, Cohen et al. 2019, ...)

Discrete-Continuous (DISCO) [this work]



✔ Equivariant
✔ Scalable

(Ocampo, Price & McEwen 2023)

Discrete-continuous (DISCO) spherical convolution

Scalable and Equivariant Spherical CNNs by Discrete-Continuous (DISCO) Convolutions
(Ocampo, Price & McEwen 2023; arXiv:2209.13603)

Follows by a careful hybrid representation of the spherical convolution:

- some components left continuous, to facilitate accurate rotational equivariance;
- while other components are discretized, to yield scalable computation.

Discrete-continuous (DISCO) spherical convolution

DISCO spherical convolution

Spherical convolution can be carefully approximated by the DISCO representation

$$(f \star \psi)(R) = \int_{\mathbb{S}^2} f(\omega)\psi(R^{-1}\omega)\mathrm{d}\omega \approx \sum_i f[\omega_i]\psi(R^{-1}\omega_i)q(\omega_i),$$

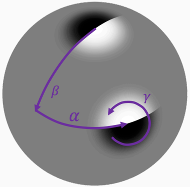
for spherical signal and filter kernel $f, \psi : \mathbb{S}^2 \to \mathbb{R}$, with spherical coordinates $\omega \in \mathbb{S}^2$, where, for now, we consider 3D rotations $R \in \mathrm{SO}(3)$.

- Appeal to sampling theorem on the sphere with quadrature weights $q : \mathbb{S}^2 \to \mathbb{R}$ (McEwen & Wiaux 2011; arXiv:1110.6298):
  - $\Rightarrow$ all information content of signal captured by samples $\{f[\omega_i]\}_i$;
  - $\Rightarrow$ continuous integral evaluated accurately by quadrature (exact for sufficient sampling).
- Filter ψ and rotation R treated continuously to avoid any discretization artefacts.

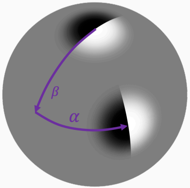
Restricting rotations to $\mathrm{SO}(3)/\mathrm{SO}(2)$

While the DISCO spherical convolution is already efficient, we seek further computational savings by restricting the space of rotations to quotient space $\mathrm{SO}(3)/\mathrm{SO}(2)$.

- **Analogous to Euclidean planar CNNs**, where filters are translated across the image but are *not* rotated in the plane.
- However, as the space $\mathrm{SO}(3)/\mathrm{SO}(2)$ is **not a group**, when restricting rotations in this manner important differences to the usual setting arise.



$$R = Z(\alpha)Y(\beta)Z(\gamma) \in \mathrm{SO}(3)$$



$$R = Z(\alpha)Y(\beta) \in \mathrm{SO}(3)/\mathrm{SO}(2) \simeq \mathbb{S}^2$$

Rotational equivariance for rotations $R \in \mathrm{SO}(3)$

DISCO spherical convolution $f \star \psi$ for rotations $Q, R \in \mathrm{SO}(3)$ satisfies SO(3) rotational equivariance:

$$((Qf) \star \psi)(R) = (Q(f \star \psi))(R).$$

Only holds since SO(3) exhibits a group structure and so $Q^{-1}R \in \mathrm{SO}(3)$.

Asymptotic rotational equivariance for rotations $R \in \mathrm{SO}(3)/\mathrm{SO}(2)$

DISCO spherical convolution $f \circledast \psi$ for rotations $Q, R \in \mathrm{SO}(3)/\mathrm{SO}(2)$ **does not satisfy $\mathrm{SO}(3)$ or $\mathrm{SO}(3)/\mathrm{SO}(2)$ rotational equivariance** (in contrast to the Euclidean setting).

But DISCO spherical convolution $f \circledast \psi$ does satisfy asymptotic SO(3) equivariance as $\beta \rightarrow 0$, where $Q = Z(\alpha)Y(\beta)Z(\gamma)$.

Asymptotic SO(3) equivariance of significant practical use since content in spherical signals often orientated and similar content appears at similar latitudes, particularly for 360° panoramic photos and video.

Computationally scalable DISCO spherical convolution

DISCO convolution affords a computationally scalable implementation.

1. Spare tensor representation.
2. Memory compression.
3. Custom sparse gradients.

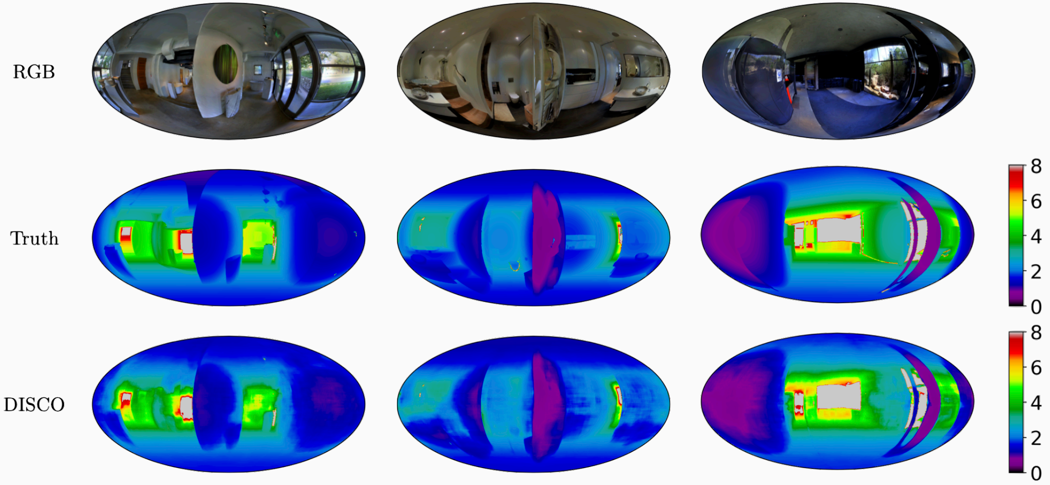
Linear scaling in number of pixels on the sphere $O(N) = O(L^2)$ for both computational cost and memory usage.

DISCO spherical CNNs exhibit

excellent rotational equivariance and are highly computationally scalable,

supporting high-resolution input/output data for dense-prediction tasks.

Depth estimation for Pano3D

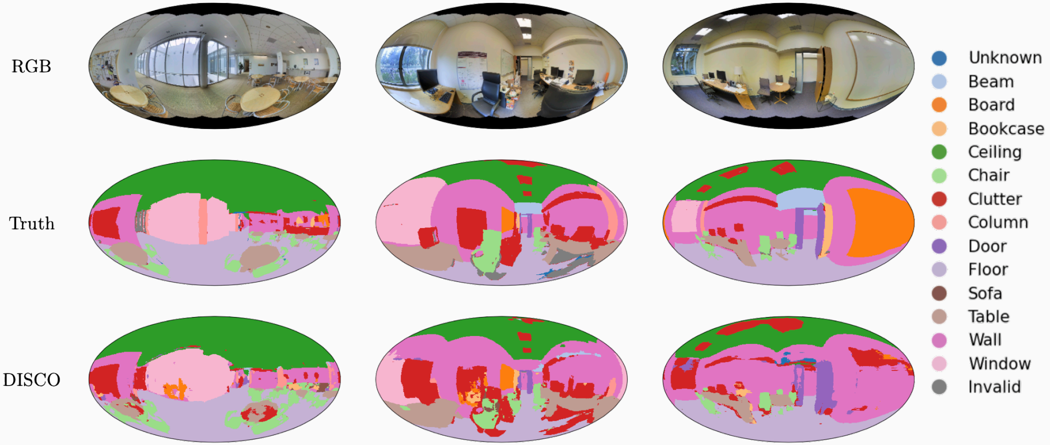


Example predictions for depth estimation of Pano3D data (depth plotted in meters).

Depth estimation for Pano3D

Model	Parameters	Depth Error Metrics				Depth Accuracy Metrics			
		*w*RMSE	*w*RMSLE	*w*AbsRel	*w*SqRel	$\delta^{\mathrm{ico}}_{1.05}$	$\delta^{\mathrm{ico}}_{1.1}$	$\delta^{\mathrm{ico}}_{1.25}$	$\delta^{\mathrm{ico}}_{1.25^2}$
Planar UNet	27M	**0.4520**	**0.1300**	0.1147	**0.0811**	36.68%	60.59%	88.31%	96.96%
DISCO-Directional (Ours)	**658k**	0.5063	0.1695	**0.1109**	0.0852	**38.32**%	**62.12**%	**88.65**%	**97.29**%

Semantic segmentation for 2D3Ds dataset



Example predictions for semantic segmentation of 2D3DS data.

Semantic segmentation for 2D3Ds dataset

Model	mIoU	mAcc
Planar UNet	35.9	50.8
UGSCNN	38.3	54.7
GaugeNet	39.4	55.9
HexRUNet	43.3	58.6
SWSCNNs	43.4	58.7
CubeNet	45.0	62.5
MöbiusConv	43.3	60.9
TangentImg	41.8	54.9
HoHoNet	43.3	53.9
DISCO-Directional-Aug (Ours)	**45.7**	**62.7**

Summary

Scalable and Equivariant Spherical CNNs by Discrete-Continuous (DISCO) Convolutions
(Ocampo, Price & McEwen 2023; arXiv:2209.13603)

Can scale rotationally equivariant spherical CNNs to high-resolution input/output data for dense-prediction tasks.

SOTA performance on all benchmark problems considered to date.

Code available on request at https://kagenova.com/products/copernicAI/ or contact jason.mcewen@kagenova.com.