# 2019sp-cs-8501-Deep2Read Scribe Notes: Spherical CNNs

Scribe: Arshdeep Sekhon

June 1, 2019

## 1 Motivation

Need for models that can analyze spherical images (for example, drones, robots, sensors,etc.). 2D CNNs/translational convolution don't work because any planar projection of a spherical signal will result in distortions. Spherical CNNs introduces spherical convolution that is rotation invariant for spherical images.

## 2 Key Challenges

- The convolution is on $S^2$ sphere, no perfectly symmetrical grids like pixel grids in images for the sphere exist: How to define rotation by a pixel?
- computational efficiency: SO(3) is a three-dimensional manifold, so a naive implementation of SO(3) correlation is O($n^6$).

## 3 Background

**The Unit Sphere $S^2$** : The set of points $x \in R^3$ with unit norm. It is a two-dimensional manifold, which can be parameterized by spherical coordinates $\alpha, \beta$.

**Spherical Filters:** $f : S^2 \rightarrow R^K$, where K is the number of channels.

**Rotations:** Any orientation can be defined by 3 elemental rotations. The set of rotations in three dimensions is called SO(3), the "special orthogonal group". parameterized by ZYZ-Euler angles.

**Rotating a point on a sphere:** If we represent points on the sphere as 3D unit vectors x, we can perform a rotation using the matrix-vector product $Rx$. The rotation group SO(3) is a three-dimensional manifold. Rotation has three degrees of freedom : $R(\alpha; \beta; \gamma) = Z(\alpha)Y(\beta)Z(\gamma) \in SO(3)$

## 4 Correlation on the sphere and rotation group

Any rotation can be defined By analogy with 2D planar CNNs:

$$f \star \psi(x) = \int f(y)\psi(x-y)dy \tag{1}$$

Define translation : $T_x^{-1}(y) = x - y$

$$f \star \psi(x) = \int f(y)\psi(T_x^{-1}(y))dy \tag{2}$$

Extending to rotations:
For first Layer:

$$f \star \psi(x) = \int_{S^2} f(y)\psi(R_x^{-1}(y))dy \tag{3}$$

After the first layer defined on SO(3):

$$f \star \psi(R) = \int_{SO(3)} f(Q)\psi(R^{-1}(Q))dQ \tag{4}$$

where 3d rotation is defined by: $R_x(t) = R_x\dot{t}$

**Spherical Rotation is equivariant to rotation** For continuous functions $f$ and $\psi$ : If rotation operator is defined as: $[L_R f](x) = f(R^{-1}x)$

$$[L_R f] \star \psi = L_R[f \star \psi] \tag{5}$$

## 5 Results

**Equivariance Error:** Equivariance error introduced because of discretization of $f$ and $\psi$. Equivariance Error $\Delta$:

$$\Delta = \frac{1}{n}\Sigma_{i=1}^{n}\frac{std(L_{Ri}\phi(f_i) - \phi(L_{Ri}f_i))}{std(\phi(f_i))} \tag{6}$$

- the approximation error grows with the resolution and the number of layers

**Rotated MNIST on sphere**

- MNIST dataset projected on the sphere
- Version 1(NR): each digit is projected on the northern hemisphere
- Version 2(R): each projected digit is additionally randomly rotated.
- trained each model on the non rotated (NR) and the rotated (R) training set and evaluated it on the non-rotated and rotated test set.
- When trained on NR and tested on R, the spherical CNN shows a slight decrease in performance compared to when trained and tested on Rotated, but still performs very well.

| | NR / NR | R / R | NR / R |
|---|---|---|---|
| planar | 0.98 | 0.23 | 0.11 |
| spherical | 0.96 | 0.95 | 0.94 |

Rotated MNIST Results

| Method | Author | RMSE |
|---|---|---|
| MLP / random CM | (a) | 5.96 |
| LGIKA(RF) | (b) | 10.82 |
| RBF kernels / random CM | (a) | 11.40 |
| RBF kernels / sorted CM | (a) | 12.59 |
| MLP / sorted CM | (a) | 16.06 |
| **Ours** | | **8.47** |

| $S^2$CNN | Layer | Bandwidth | Features |
|---|---|---|---|
| | Input | | 5 |
| | ResBlock | 10 | 20 |
| | ResBlock | 8 | 40 |
| | ResBlock | 6 | 60 |
| | ResBlock | 4 | 80 |
| | ResBlock | 2 | 160 |
| DeepSet | Layer | Input/Hidden | |
| | $\phi$ (MLP) | 160/150 | |
| | $\psi$ (MLP) | 100/50 | |

QM7 task results

### Prediction of Atomization Energies from Molecular Geometry

- QM7 task
- the atomization energy of molecules to be predicted from geometry and charges