Spherical CNNS

Liangjie Cao

July 8, 2018

1. Correlation on the Sphere and Rotation Group

In order to define the spherical correlation, the authors need to know not only how to rotate points $x \in S^2$ but also how to rotate filters (*i.e.* functions) on the sphere. To this end, they introduce the rotation operator L_R that takes a function f and produces a rotated function $L_R f$ by composing f with the rotation R^{-1}:

$$[L_R f](x) = f(R^{-1}x) \tag{1}$$

The inner product on the vector space of spherical signals is defined as:

$$< \psi, f > = \int_{S^2} \sum_{k=1}^{K} \psi_k(x) f_k(x) dx \tag{2}$$

The integration measure dx denotes the standard rotation invariant integration measure on the sphere, which can be expressed as $d\alpha\, R\, sin(\beta) d\beta / 4\pi$ in R spherical coordinates. The invariance of the measure ensures that $\int_{S^2} f(Rx) dx = \int_{S^2} f(x) dx$, for any rotation $R \in SO(3)$. That is, the volume under a spherical height map does not change when rotated. Using this fact, they can show that $L_{R^{-1}}$ is adjoint to L_R, which implies that L_R is unitary:

$$< L_R \psi, f > = \int_{S^2} \sum_{k=1}^{K} \psi_k(R^{-1}x) f_k(x) dx \tag{3}$$

Then we know that this two function is equal. With these ingredients in place, we are now ready to state mathematically what was stated in words before. For spherical signals f and ψ, the authors define the correlation as:

$$[\psi \star f] = < L_R \psi, f > = \int_{S^2} \sum_{k=1}^{K} \psi_k(R^{-1}x) f_k(x) dx \tag{4}$$

As mentioned before, the output of the spherical correlation is a function on $SO(3)$. This is perhaps somewhat counterintuitive, and indeed the conventional definition of spherical convolution gives as output a function on the sphere. However, the conventional definition effectively restricts the filter to be circularly symmetric about the Z axis, which would greatly limit the expressive capacity of the network.

The authors defined the rotation operator L_R for spherical signals (Eq. 1), and used it to define spherical cross-correlation (Eq. 1). To define the $SO(3)$ correlation, they need to generalize the rotation operator so that it can act on signals defined on $SO(3)$. As we will show, naively reusing Eq. 1 is the way to go. That is, for $f : SO(3) \to \mathbb{R}^K$, and $R, Q \in SO(3)$:

$$[L_R f](Q) = f(R^{-1}Q) \tag{5}$$

Note that while the argument $R^{-1}x$ in Eq. 1 denotes the rotation of $x \in S$ by $R \in SO(3)$, the analogous term $R^1 Q$ in Eq. 1 denotes to the composition of rotations (*i.e.* matrix multiplication).

Using the same analogy as before, they can define the correlation of two signals on the rotation group, $f, \psi : SO(3) \to \mathbb{R}^K$, as follows:

$$[\psi \star f](R) = \int_{SO(3)} \sum_{k=1}^{K} \psi_k(R^{-1}Q) f_k(Q) dQ \tag{6}$$

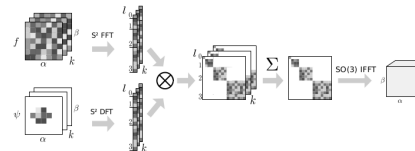The derivation is valid for spherical correlation as well as rotation group correlation.

Figure 1. Spherical correlation in the spectrum. The signal f and the locally-supported filter ψ are Fourier transformed, block-wise tensored, summed over input channels, and finally inverse transformed. Note that because the filter is locally supported, it is faster to use a matrix multiplication (DFT) than an FFT algorithm for it. They parameterize the sphere using spherical coordinates α, β, and $SO(3)$ with ZYZ-Euler angles α, β, γ

2. Fast Spherical Correlation with G-FFT

For functions on the sphere and rotation group, there is an analogous transform, which they will refer to as the generalized Fourier transform (GFT) and a corresponding fast algorithm (GFFT). This transform finds it roots in the representation theory of groups, but due to space constraints we will not go into details here and instead refer the interested reader to [3] and [1]. This process can be seen in Fig. 1.

The inverse $SO(3)$ Fourier transform is defined as:

$$f(R) = \sum_{i=0}^{b}(2l+1)\sum_{m=-l}^{l}\sum_{n=-l}^{l}\hat{f_{mn}^{l}}U_{mn}^{l}(R) \quad (7)$$

and similarly for S^2. The maximum frequency b is known as the bandwidth, and is related to the resolution of the spatial grid [2].

The algorithm for the S^2-FFTs is very similar, only in this case their FFT over the α axis only, and do a linear contraction with precomputed Legendre functions over the β axis.

Their code is available at `https://github.com/jonas-koehler/s2cnn`.

References

[1] G. B. Folland. *A Course in Abstract Harmonic Analysis*. Crc Press, 1995. 2

[2] P. J. Kostelec and D. N. Rockmore. Soft: So(3) fourier transforms. URL `https://colah.github.io/posts/2014-12-Groups-Convolution`, 2007. 2

[3] Sugiura and Mitsuo. *Unitary representations and harmonic analysis*. Kodansha, 1975. 2