Regular Group Convolution

Let f an input signal, $f: \mathbb{R}^n \longrightarrow \mathbb{R}$
Let k a filter, $k: \mathbb{R}^n \longrightarrow \mathbb{R}$. The <u>convolution</u> of f and k is

$$[k * f](x) = \int_{\mathbb{R}^n} k(\tilde{x} - x) f(\tilde{x}) \, d\tilde{x}$$

If we consider the discrete setting

$$f: \mathbb{Z}^n \longrightarrow \mathbb{R} \quad , \quad k: \mathbb{Z}^n \longrightarrow \mathbb{K}$$

and

$$[k * f](x) = \sum_{\tilde{x} \in \mathbb{Z}^n} k(\tilde{x} - x) f(\tilde{x})$$

Visually the convolution is the scalar product between k and all the possible translations of f.
Since k is <u>stationary</u> then the <u>same</u> kernel is applied on any translation of f. This makes the convolution <u>translation equivariant</u>, ie,

$$[k * \mathcal{L}_t f](x) = \int_{\mathbb{R}^n} k(\tilde{x} - x)\, f(\tilde{x} - t)\, d\tilde{x} = (\Delta)$$

substituting $\tilde{x}$ with $\hat{x} + t$

$$(\Delta) = \int_{\mathbb{R}^n} k(\hat{x} + t - x)\, f(\hat{x})\, d\hat{x} = \int_{\mathbb{R}^n} k(\hat{x} - (x-t))\, f(\hat{x})\, d\hat{x} =$$

$$= [k * f](x-t) = \mathcal{L}_t [k * f](x)$$

We prove that the standard convolution is equivariant respect to transl. of the input.

Now we are ready to define the G_convolution

<u>Definition</u>: let G a group, $f: G \longrightarrow \mathbb{R}$ and $k : G \longrightarrow \mathbb{R}$, we define G_conv as

$$[k *_G f](g) = \int_G k(g^{-1}\tilde{g})\, f(\tilde{g})\, d\tilde{g} =$$

$$= \int_G \mathcal{L}_g k(\tilde{g})\, f(\tilde{g})\, d\tilde{g}$$

<u>Remark</u> : if $G = (\mathbb{R}^m, +) \Rightarrow$ standard conv.

The G-conv is equivariant w.r.t. the action of G

$$[K * \mathcal{L}_{h\in G} f](g) = \int_G K(g^{-1}\tilde{g})\, f(h^{-1}\tilde{g})\, d\tilde{g} = \{\text{subst. } \tilde{g} \leftarrow h\tilde{g}\} =$$

$[d\tilde{g} = d(h\tilde{g})$ why?$]$

$$= \int_G K(g^{-1}\tilde{g})\, f(\cancel{h^{-1}}\cancel{h}\tilde{g})\, d\tilde{g} = \int_G K(\overbrace{(g^{-1}h)}^{\in G}\tilde{g})\, f(\tilde{g})\, d\tilde{g} =$$

$$= [K * f](h^{-1}g) = \mathcal{L}_{h\in G}[K * f](g)$$

The objective is gaining equivariance wrt a group bigger than $N = (\mathbb{R}^m, +)$, for example adding rotation. In this case $G = (\mathbb{R}^m, +) \rtimes H$ where H is the "rotational" part i.e. $SO(n)$.

<u>Remark</u> : The following construction of semidirect product are equivalent.

(i) [internal] : Let $N \trianglelefteq G$ and $H \leq G$ such that $G = NH$ and $H \cap N = 1_G$

Then the map $\Phi : H \longrightarrow Aut(N)$ is
$$h \longmapsto h^{-1} n h$$
a group homomorfism with $Ker(\Phi) = C_H(N)$

In this case G is the semidirect product of N and H via Φ.

(ii) [external]: N, H groups and $\Phi : H \longrightarrow Aut(N)$ homomorfism. $\forall x \in H$ and $\forall a \in N$ we write $a^{\phi(x)}$ for $\phi(x)(a)$. On the <u>set</u> $N \times H$ we define the following operation
$$(a, x)(b, y) = (a\, b^{\phi(x)^{-1}}, xy)$$
with this operation $N \times H$ is a group called semidirect product of N and H, ie, $G = N \rtimes H$. ☒

In special case such as $G = (\mathbb{R}^n, +) \rtimes H$ we can factorize the action in the convolution operation. This will be <u>critical</u> in applications. Let $N = (\mathbb{R}^n, +)$, H group, $G = N \rtimes H$. We can think f and k defined on G as function that depend on $t \in N$ and $h \in H$ separately. Then it holds

$$[K \underset{G}{*} f](t,h) = \int_N \int_H \mathcal{L}_t \mathcal{L}_h \, K(\tilde{t},\tilde{h}) \, f(\tilde{t},\tilde{h}) \frac{1}{|h|} \, d\tilde{t} d\tilde{h} =$$

$$= \int_N \int_H \mathcal{L}_t \, K(t, h^{-1}\tilde{h}) \, f(\tilde{t},\tilde{h}) \frac{1}{|h|} \, d\tilde{t} \, d\tilde{h} =$$

$$= \int_N \int_H K(h^{-1}(\tilde{t}-t), h^{-1}\tilde{h}) \, f(\tilde{t},\tilde{h}) \frac{1}{|h|} \, d\tilde{t} d\tilde{h} =$$

where $\frac{1}{|h|}$ is a normalization factor that adj volumes if the H group change it. In particular is the determinant of the linear representation of $h \in H$ in GL. In case of rotation is 1.

<u>Remark</u>: In discrete regime

$$[K \underset{G}{*} f](t,h) = \sum_{t \in N} \sum_{h \in H} K(h^{-1}(\tilde{t}-t), h^{-1}\tilde{h}) f(\tilde{t},\tilde{h}) \frac{1}{|h|}$$

<u>Remark</u>: Since in the application we start from $\mathbb{R}^2$, ie an image with a fix orientation, the first action is called lifting and follows the following.

$$[K \overset{lift}{\underset{G}{*}} f](t,h) = \int_N K(h^{-1}(\tilde{t}-t))\, f(\tilde{t})\, \frac{1}{|h|}\, d\tilde{t}$$

The output is in $G = N \rtimes H$.

STEERABLE CNNS

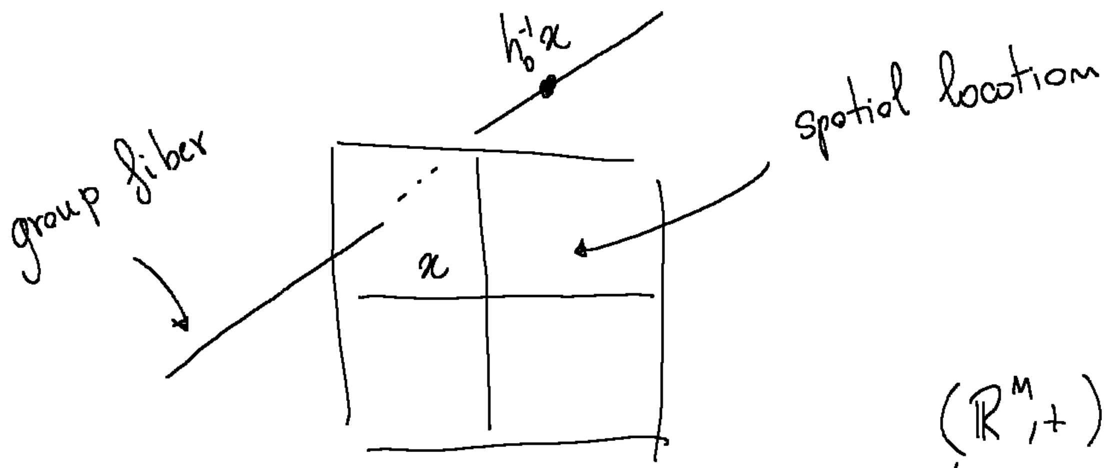
When we talk about $SO(n)$ as group H we actually consider a discrete subgroup of $SO(n)$ because continuos group are intractable using the previous construction. Steerable CNNS overcome this limitation.

Instead of computing correlation locally for each transformed version of the kernel, we compute the correlation with a carefully chosen equivariant basis at each spatial location and use these to compute a specific response. Is the same as Fourier coefficient for Fourier transform.

Feature Field: In G-CNNS a layer ℓ_G that is equiv. respect to G maps the input space $\mathbb{R}^n \rtimes H$ (or $\mathbb{R}^N$) to $\mathbb{R}$ => the features can be see as scalar field with an added dimention respect to CNN.

Alternatively ℓ_G can be seen as $\ell_G : \mathbb{R}^n \longrightarrow \mathbb{R}^{|H|}$ where for each spatial location $x \in \mathbb{R}^N$ we have a feature vector $f(x) \in \mathbb{R}^{|H|}$.

These feoutures vector are called <u>group fiber</u>



For a given $g = (t,h) \in G = N \rtimes H$ the feature field of the layer l_G transforms as

$(\mathbb{R}^M, +)$ ↙ N

$$[\mathcal{L}_G \, l_G](x) = \rho_{reg}^H(h) \, l_G(g^{-1}x)$$

where $\rho^H_{reg}(h)$ is the regular representation <u>matrix</u> of $h \in H$. These field are called regular fied

↳ [dimension $|H| \times |H|$!!]

Thus the action of $g = (t,h)$ on the regular field can be visualized as <u>relocating</u> the group fibers using the action of the <u>entire group</u> G and <u>transforming</u> the group fiber themself using <u>only the action of H</u>.

The steerable CNNs improve this, introducing a steerable feature field $f : \mathbb{R}^n \longrightarrow \mathbb{R}^c$ where each x is mapped in a feature vector in $\mathbb{R}^c$

A steerable G-equiv. layer transform $\hat{\ell}_G$ as

$$[\mathcal{L}_g \hat{\ell}_G](x) = \rho_H(h)\, \hat{\ell}_G(g^{-1}x)$$

The difference is that here we used a represent. of H that is not the regular one. Thus

$$\rho_H : \begin{aligned} H &\longrightarrow GL(\mathbb{R}^c) \\ h &\longmapsto \rho_H(h) \end{aligned}$$

These representation are often modelled with irriducible representation. This make possible to treat infinite groups. Indeed if $|H| = \infty$ then "$\rho^H_{reg}(h) \in M_{\infty \times \infty}$" while $\rho_H(h) \in M_{c \times c}$.

Before going deep on the steerable GCNNs we need some background on <u>Representation Theory</u>

<u>Def</u>: A linear group representation ρ of a group G on a vector space V is a group homomorfism

$$\rho : G \longrightarrow GL(V)$$

st $\rho(g_1 g_2) = \rho(g_1)\rho(g_2)$

<u>Ex</u>: $\rho : G \longrightarrow GL(\mathbb{R})$ is the <u>trivial represent.</u>
$g \longmapsto 1$

<u>Ex</u>: $\rho : SO(2) \longrightarrow GL(\mathbb{R}^2)$

$$r_\theta \longmapsto \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

is a representation of $SO(2)$ on $\mathbb{R}^2$

Two representations ρ e ρ' on V are <u>equivalent</u> iff are related by a change of bases $Q \in GL(V)$

$$\forall g \in G \quad \rho'(g) = Q\rho(g)Q^{-1}$$

Def: Let $\rho_1 : G \longrightarrow GL(V_1)$ and $\rho_2 : G \longrightarrow GL(V_2)$

We define $\rho_1 \oplus \rho_2 : G \longrightarrow GL(V_1 \oplus V_2)$

$$(\rho_1 \oplus \rho_2)(g) = \begin{bmatrix} \rho_1(g) & 0 \\ 0 & \rho_2(g) \end{bmatrix}$$

Ex: $\rho : G \longrightarrow GL(V)$ gr. repr. and $V_1 \subset V$ st $\rho(g)(V_1) \subseteq V_1$. Then V_1 is invoriant subspoa for the oction of ρ. Let consider V_2 st $V_1 \oplus V_2 = V$. If we choose a bosis of consistent with this decompositione them

$$\rho(g) = \begin{bmatrix} \rho_1(g) & 0 \\ 0 & \rho_2(g) \end{bmatrix}$$

and ρ_1, ρ_2 are the restriction of ρ to V_1 and V_2.

Def: A represeutotiom is irriducible (irreps) if it does not contoim mon-triviol invoriom subspoa. Instead of ρ we use ψ to be more clior.

Thm: Any unitary representation $\rho : G \longrightarrow V$ over $\mathbb{R}^N$ is a direct sum of irreps. In particular any linear representation of $\rho : G \longrightarrow \mathbb{R}^n$ of a compact group G can be decomposed as

$$\rho(g) = Q \left[\bigoplus_{i \in I} \psi_i(g) \right] Q^{-1}$$

set of ψ in ρ

Ex: Irreps of $SO(2)$

(i) $\psi_0^{SO(2)}(r_\vartheta) = 1 \rightsquigarrow$ Trivial

(ii) $\psi_k^{SO(2)}(r_\vartheta) = \begin{bmatrix} \cos(k\vartheta) & -\sin(k\vartheta) \\ \sin(k\vartheta) & \cos(k\vartheta) \end{bmatrix} = \psi(k\vartheta)$, $k \in \mathbb{N}$

Regular Representations: G finite group and let $\mathbb{R}^{|G|}$ be a vector sp. Each basis vector e_g of $\mathbb{R}^{|G|}$ is associated to a $g \in G$. The regular repr.

$$\rho_{reg}^G : G \longrightarrow GL(\mathbb{R}^{|G|})$$
$$\tilde{g} \longmapsto \rho_{reg}^G(\tilde{g}) \rightsquigarrow \text{permutation matrix}$$

The action on e_g of $\rho_{reg}^G(\tilde{g})$ is

$$\rho_{reg}^G(\tilde{g}) \cdot e_g = e_{g\tilde{g}}$$

ie $\rho_{reg}^G(\tilde{g})$ permutes the basis.

Thm (Peter_Weyl): G group compact and ρ repr. on V. Let $L_2(G)$ the vector space of squared integrable function over G.

Consider the regular repr. of G on $L^2(G)$

$$f \longmapsto \rho(g) : [\mathcal{L}_g f](h) = f(g^{-1}h)$$

This representation decomposes as sum of irreps.

In particular the matrix coeff of irreps of G span $L_2(G)$.

The set $\{ \sqrt{d_\psi} \, [\psi(g)]_{ij} \mid \psi \in \hat{G} \, , \, 1 \leq i,j \leq d_\psi \}$

| END REPR. Theory |

TO BE FINISH