HARMONIC NETWORKS:

DEEP TRANSLATION AND ROTATION EQUIVARIANCE

Daniel E. Worrall, Stephan J. Garbin,
Daniyar Turmukhambetov, and Gabriel J. Brostow

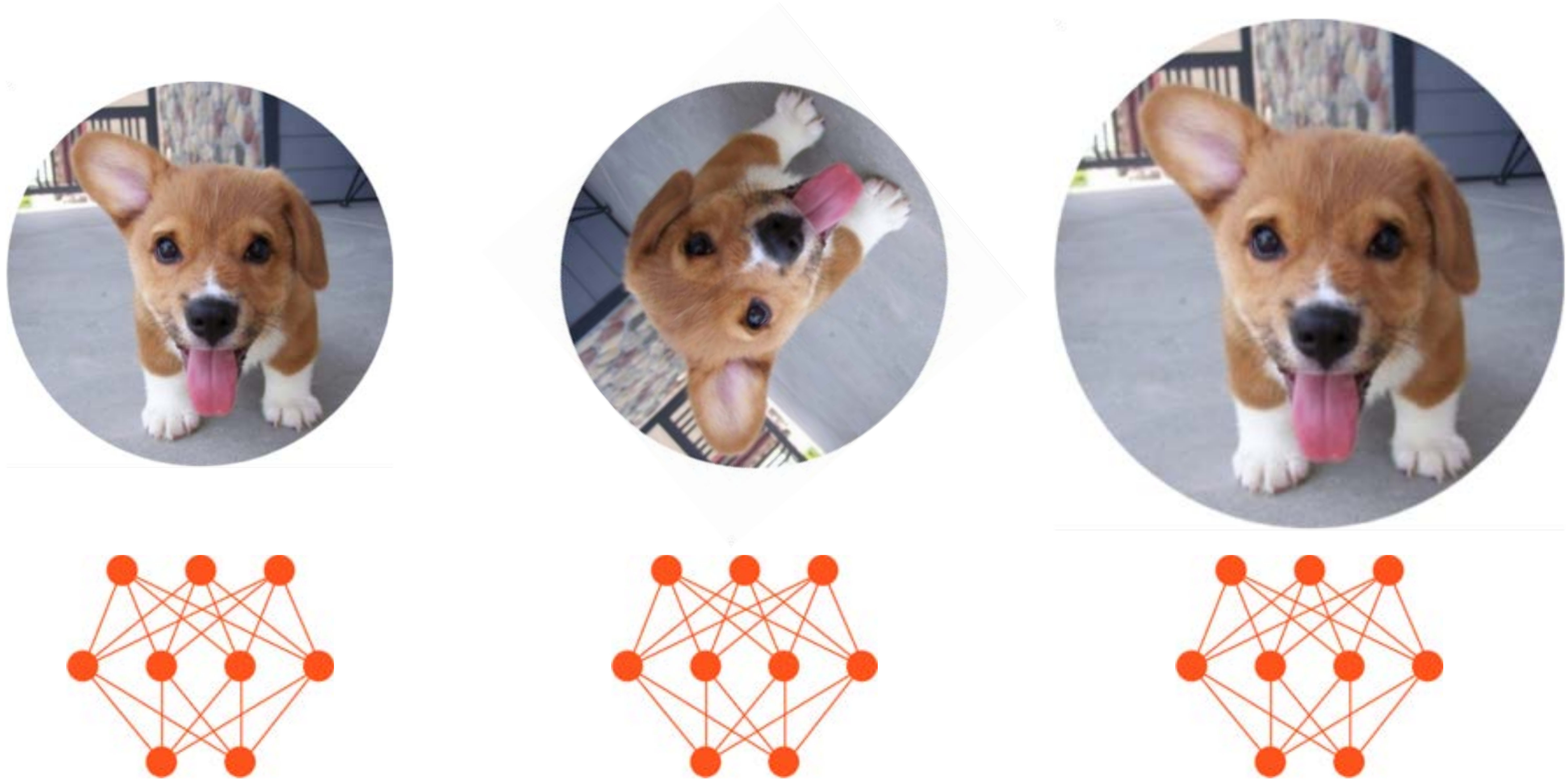
OUR GROUP

OUTLINE

- Brief intro: Invariance and Equivariance
- What has been done for rotations?
- Harmonic Networks and the Fourier Shift Theorem
- Appendix: Further theory and future works

INVARIANCE

- Many tasks **in**variant to input transformations

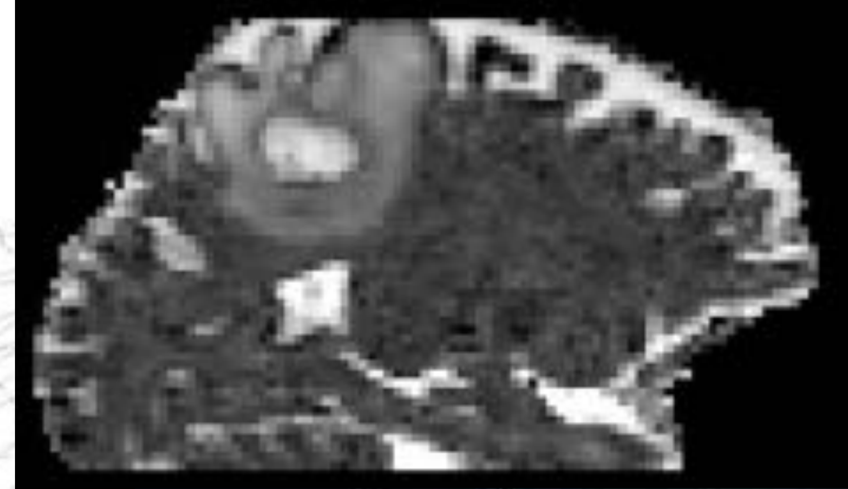


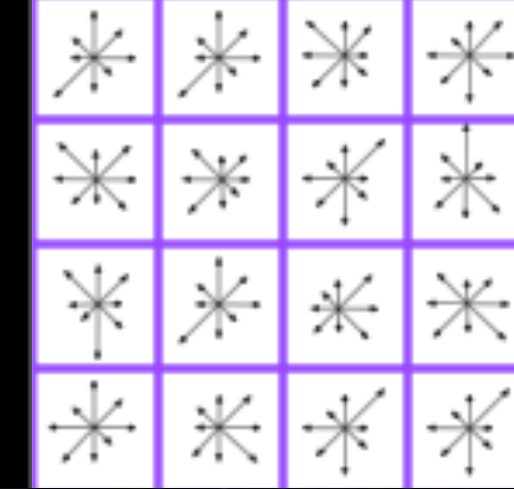
Doggie

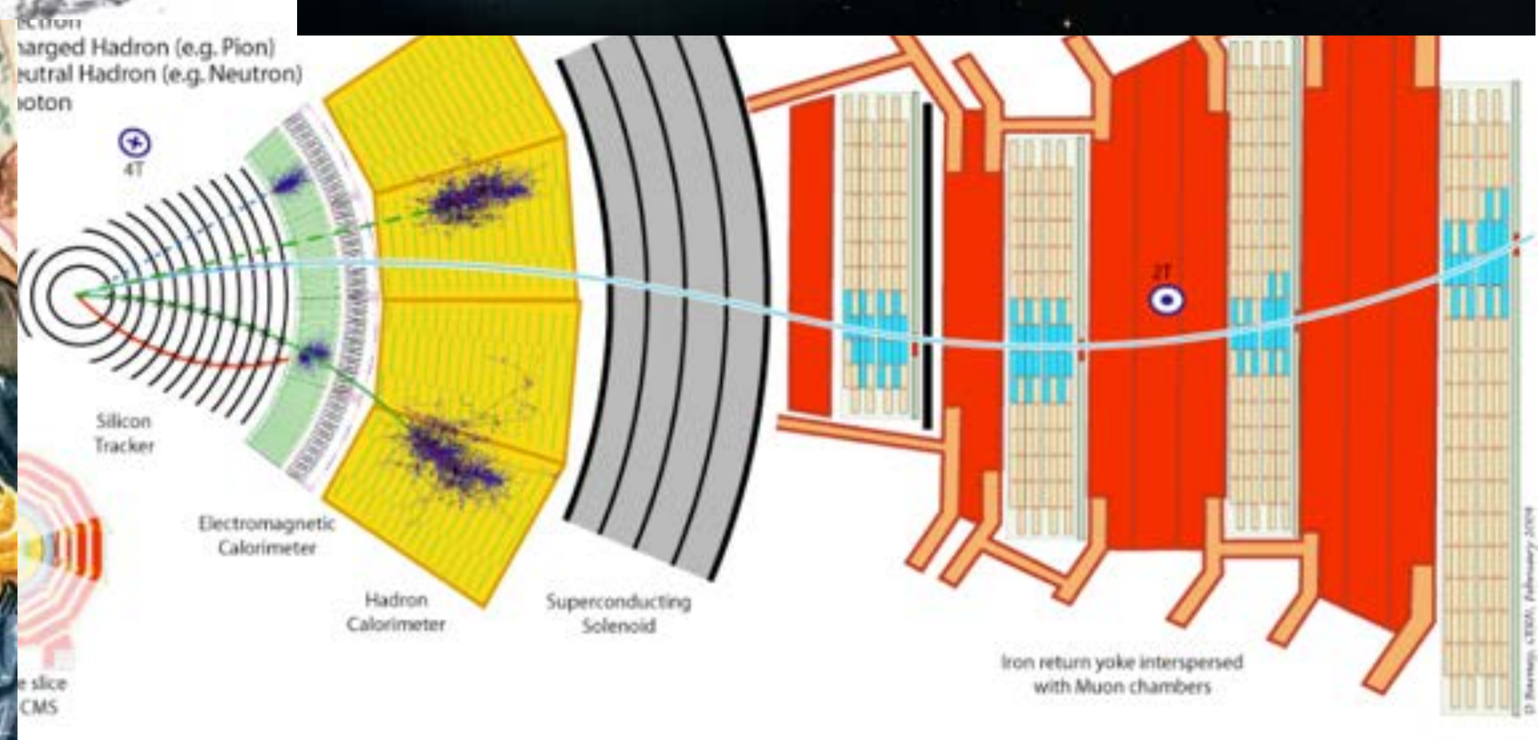
Doggie

Doggie

INVARIANCE






harged Hadron (e.g. Pion)
eutral Hadron (e.g. Neutron)
hoton
4T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
2T
Iron return yoke interspersed with Muon chambers

INVARIANCE

WARNING: DIFFERENT NOTATION TO PAPER

Mapping independent of T_g, for all T_g

$$f(\mathbf{I}) = f(\mathcal{T}_g[\mathbf{I}])$$

Notational aside: (in our case)

$$\mathcal{T}_g[\mathbf{I}](\mathbf{x}) = \mathbf{I}(\mathbf{M}_g\mathbf{x} + \mathbf{b_g})$$

e.g. translation

$$\mathcal{T}_g[\mathbf{I}](\mathbf{x}) = \mathbf{I}(\mathbf{x} + \mathbf{b}_g)$$

INVARIANCE: EXAMPLES

Fourier spectrum modulus: translations

$$\left|\int_{-\infty}^{\infty} I(x) e^{-\iota\omega x}\,\mathrm{d}x\right| = \left|\int_{-\infty}^{\infty} I(x-b) e^{-\iota\omega x}\,\mathrm{d}x\right|$$

Eigenspectrum: orthogonal transformations

$$\mathrm{spectrum}[\mathbf{M}] = \mathrm{spectrum}[\mathbf{Q}\mathbf{M}\mathbf{Q}^\top] \qquad \mathbf{Q}\mathbf{Q}^\top = \mathbf{Q}^\top\mathbf{Q} = \mathbf{I}$$

Variance: translations

$$\mathrm{Var}[p] = \mathrm{Var}[\mathcal{T}_g p]$$

Pooling: smooth invertible mappings (typ. diffeomorphisms)

$$\int_{\mathbf{x}\in\Omega} I(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbf{x}\in\Omega} I(\mathcal{T}_g\mathbf{x})\,\mathrm{d}\mathbf{x} \qquad \max_{\mathbf{x}\in\Omega} I(\mathbf{x}) = \max_{\mathbf{x}\in\Omega} I(\mathcal{T}_g\mathbf{x})$$

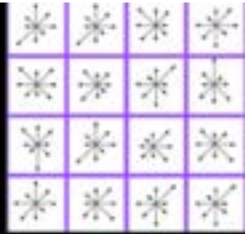
EQUIVARIANCE

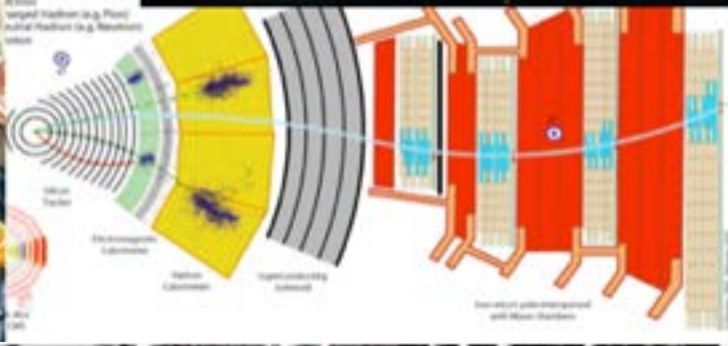
“The man swallowed the chicken”

versus

“The chicken swallowed the man”

INVARIANCE





EQUIVARIANCE: GENERAL DEFINITION

Mapping preserves algebraic structure of transformation

$$\underbrace{\mathcal{S}_g [f(\mathbf{I})] = f(\mathcal{T}_g}_{\text{Different \textit{representations} of same transformation}}[\mathbf{I}])$$

Different *representations* of
same transformation

Invariance is special case of equivariance

$$\mathcal{S}_g = \mathrm{Id}$$

EQUIVARIANCE: EXAMPLES

Fourier spectrum phase: translations

$$e^{-\iota\omega b}\int_{-\infty}^{\infty} I(x)e^{-\iota\omega x}\,\mathrm{d}x = \int_{-\infty}^{\infty} I(x-b)e^{-\iota\omega x}\,\mathrm{d}x$$

Convolution (and correlation)

$$[f * g](x-b) = [T_b[f] * g](x)$$

Mean: translations

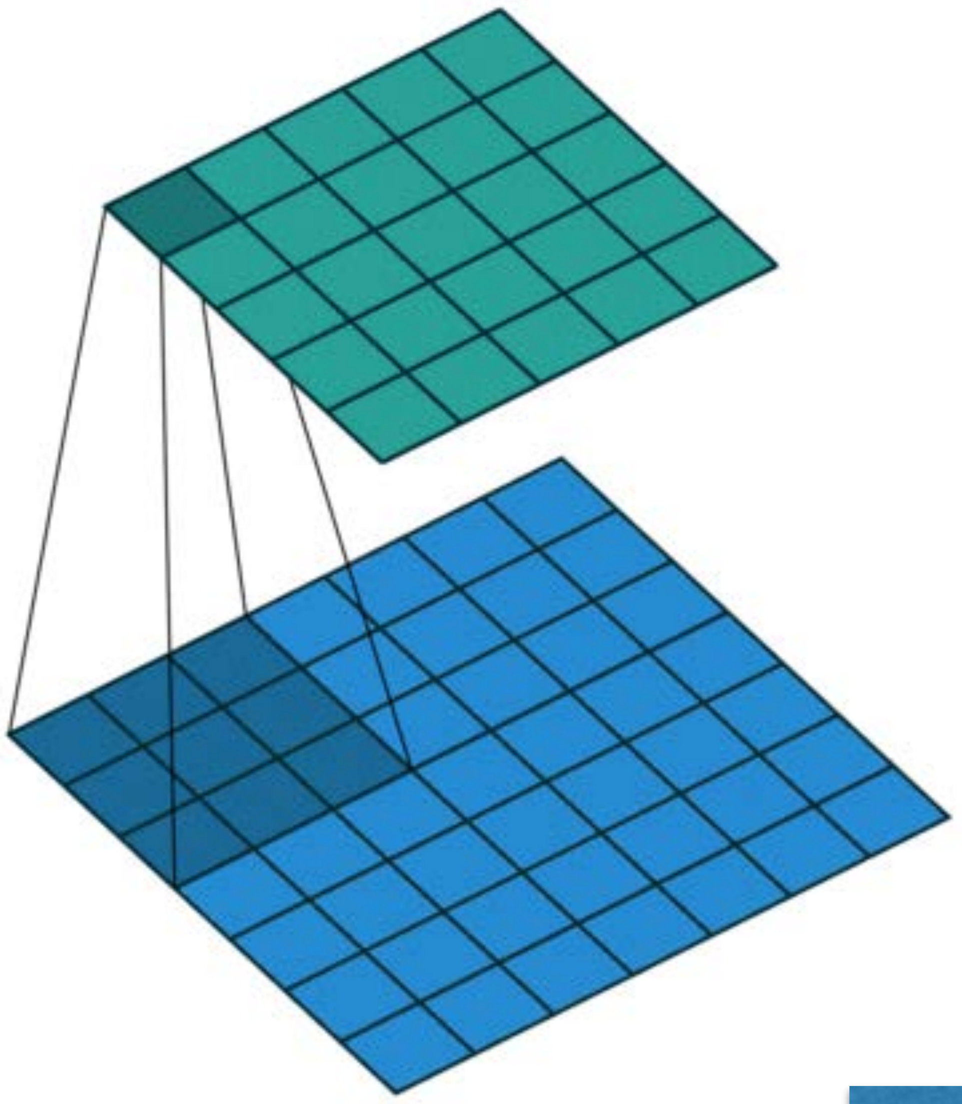
$$a\mathbb{E}_p[x] + b = \mathbb{E}_p[ax+b]$$

Fourier transform: differentiation

(and pretty much any other linear transformation)

$$|\iota\omega|^n \hat{h}(\omega) = \widehat{\partial_x^n h}(\omega)$$

TRANSLATION **EQUI**VARIANCE OF Z-CORRELATION

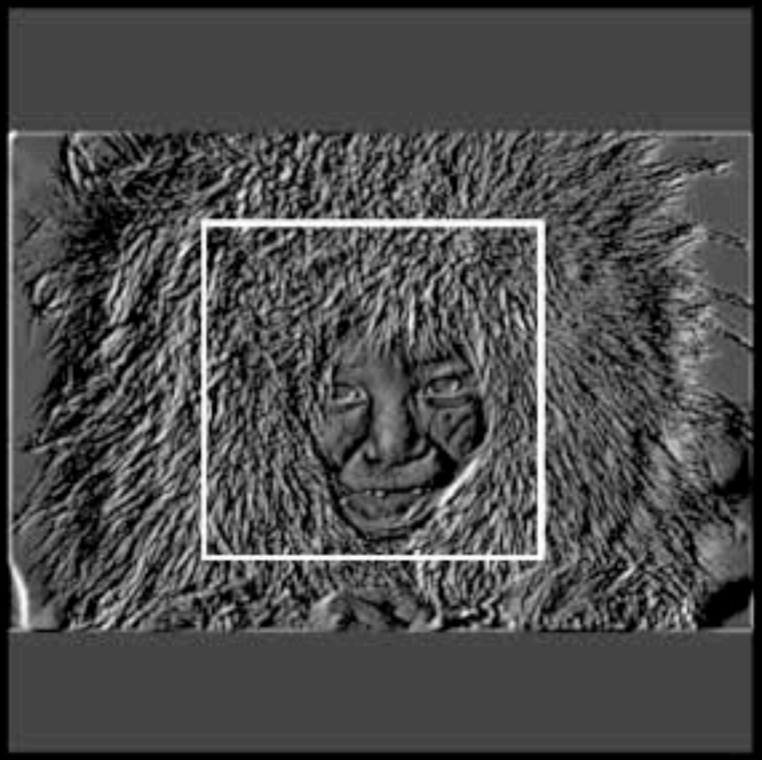


$$
\begin{aligned}
[w \star \mathcal{T}_b f](\mathbf{x}) &= \sum_{\mathbf{y} \in \mathbf{X}} w(\mathbf{y} - \mathbf{x}) \mathcal{T}_b f(\mathbf{y}) \\
&= \sum_{\mathbf{y} \in \mathbf{X}} w(\mathbf{y} - \mathbf{x}) f(\mathbf{y} - \mathbf{b}) \\
&= \sum_{\mathbf{y}' \in \mathbf{X}} w(\mathbf{y}' + \mathbf{b} - \mathbf{x}) f(\mathbf{y}') \\
&= [w \star f](\mathbf{x} - \mathbf{b})
\end{aligned}
$$

Boundaries imply local equivariance

Source: https://github.com/vdumoulin/conv_arithmetic

LACK OF ROTATION **EQUI**VARIANCE:



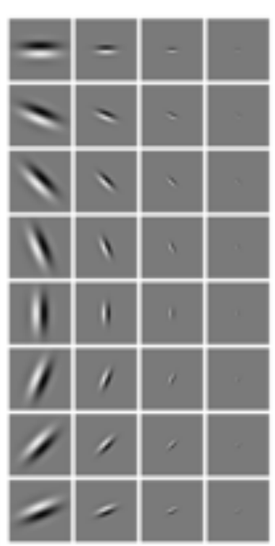
ROTATION **EQUI**VARIANCE METHODS

Rotate in network

Rotate filters

Store full response

Roto-Translation Scattering
(Oyallon et al., 2014)
Group Equivariant CNNs
(Cohen et al., 2016)
Oriented Response Networks
(Zhou et al., 2017)
Steerable CNNs
(Cohen et al., 2017)

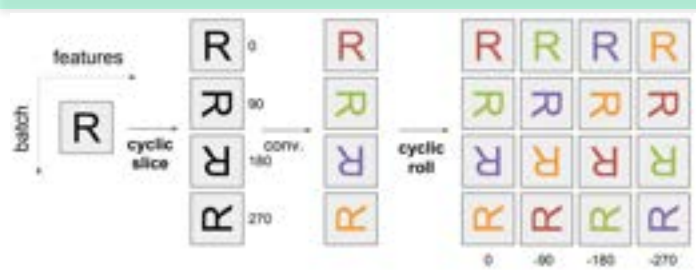


Store max/mean response

Roto-Translation Invariant Scattering
(Sifre et al., 2013)
Rotation Equivariant Vector Fields
(Marcos et al., 2017)

Rotate activations

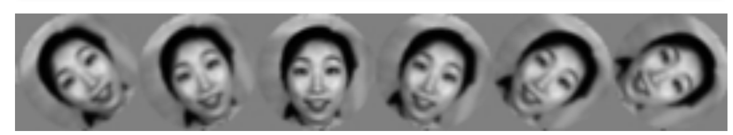
Exploiting Cyclic Symmetry
(Dieleman et al., 2016)



Rotate input

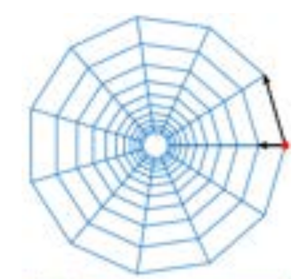
Brute-force

Rotation-Invariant Neoperceptron
(Fasel et al., 2006)
TI-Pooling
(Laptev et al., 2016)



Log-polar trans.

Warped Convolutions
(*Henriques et al., 2017)*
Polar Transformer Nets
(Esteves et al., 2017)



ALL DISCRETE ROTATIONS

"Do you want the wrong answer to the right question or the right answer to the wrong question? I think you want the former."

DAVID BLEI

PART II:
HARMONIC NETWORKS

THE FOURIER SHIFT THEOREM

Interestingly this generalises beyond 1D shifts to any transformation if considering groups

$$\int_{-\pi}^{\pi} I(\phi - \theta) e^{-\iota m \phi} \, \mathrm{d}\phi = \int_{-\pi-\theta}^{\pi-\theta} I(\phi') e^{-\iota m (\phi' + \theta)} \, \mathrm{d}\phi'$$

$$m \in \mathbb{Z} \qquad = \underline{e^{-\iota m \theta}} \underline{\int_{-\pi}^{\pi} I(\phi') \boxed{e^{-\iota m \phi'}} \, \mathrm{d}\phi'}$$

Equivariant Invariant

Constraint on angular component of weights

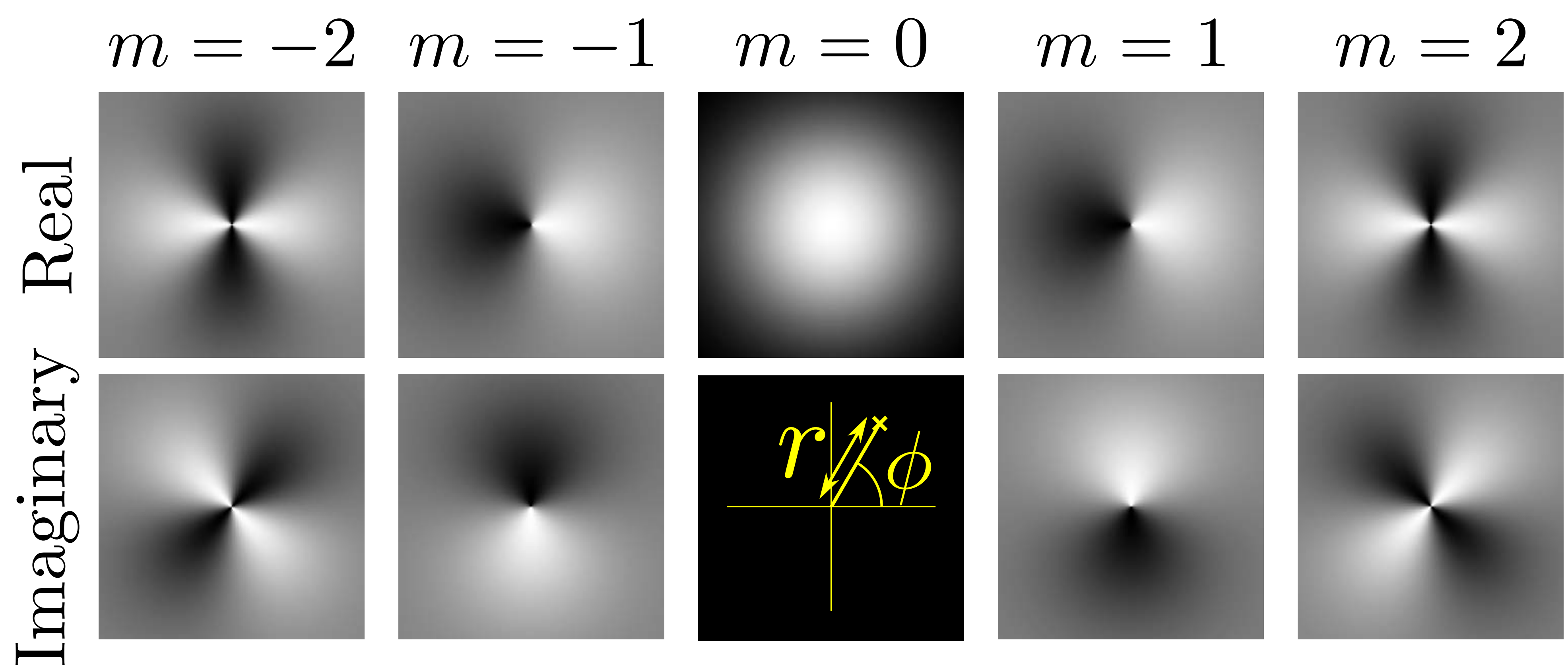
CIRCULAR HARMONICS

Note: the sign flip in the angular component

Live in orthogonal spaces, so don't interact

$$W_m(r;\phi) = \underline{R(r)}e^{\iota(\underline{m\phi}+\underline{\beta})}$$

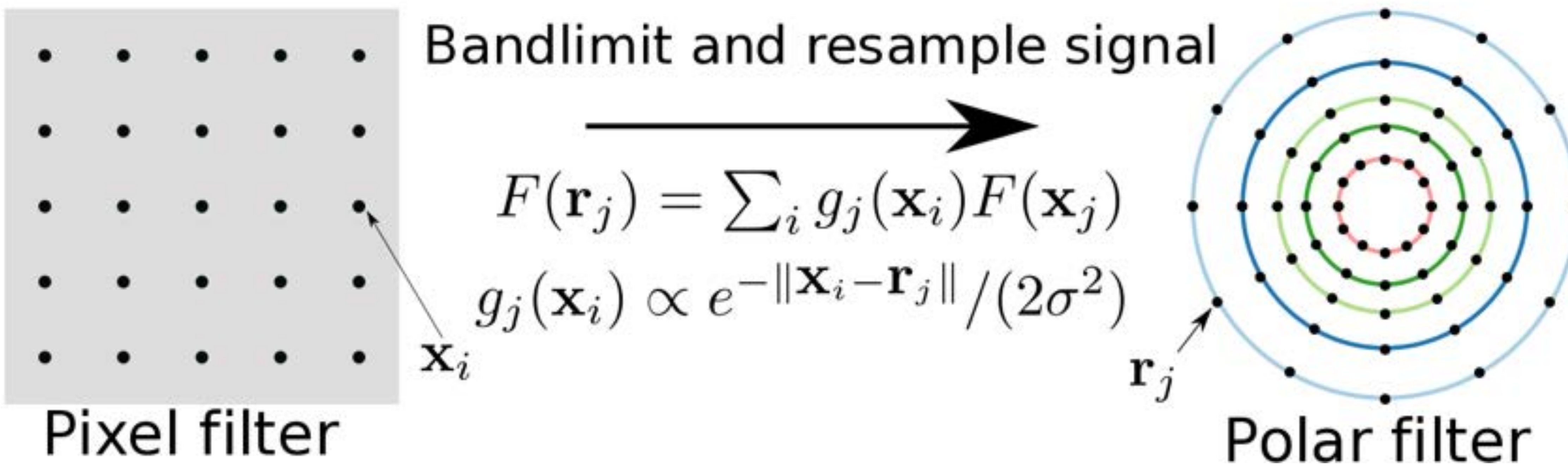
Free Constrained



Example with Gaussian radial component

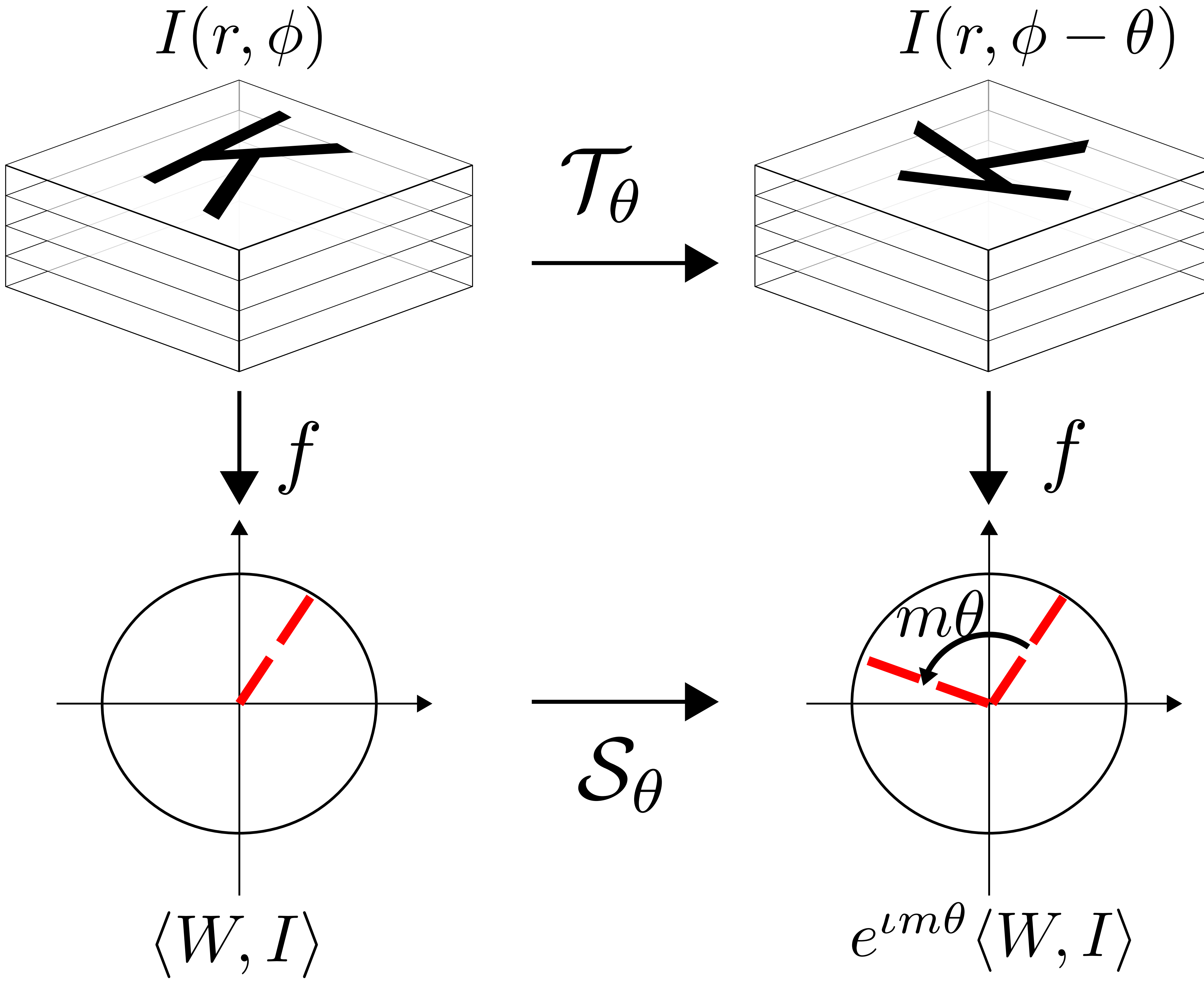
CIRCULAR HARMONIC WEIGHTS: IN PRACTICAL TERMS

Circular harmonics are defined in polar coordinates, so **resample** with Gaussian **anti-aliasing filter**.



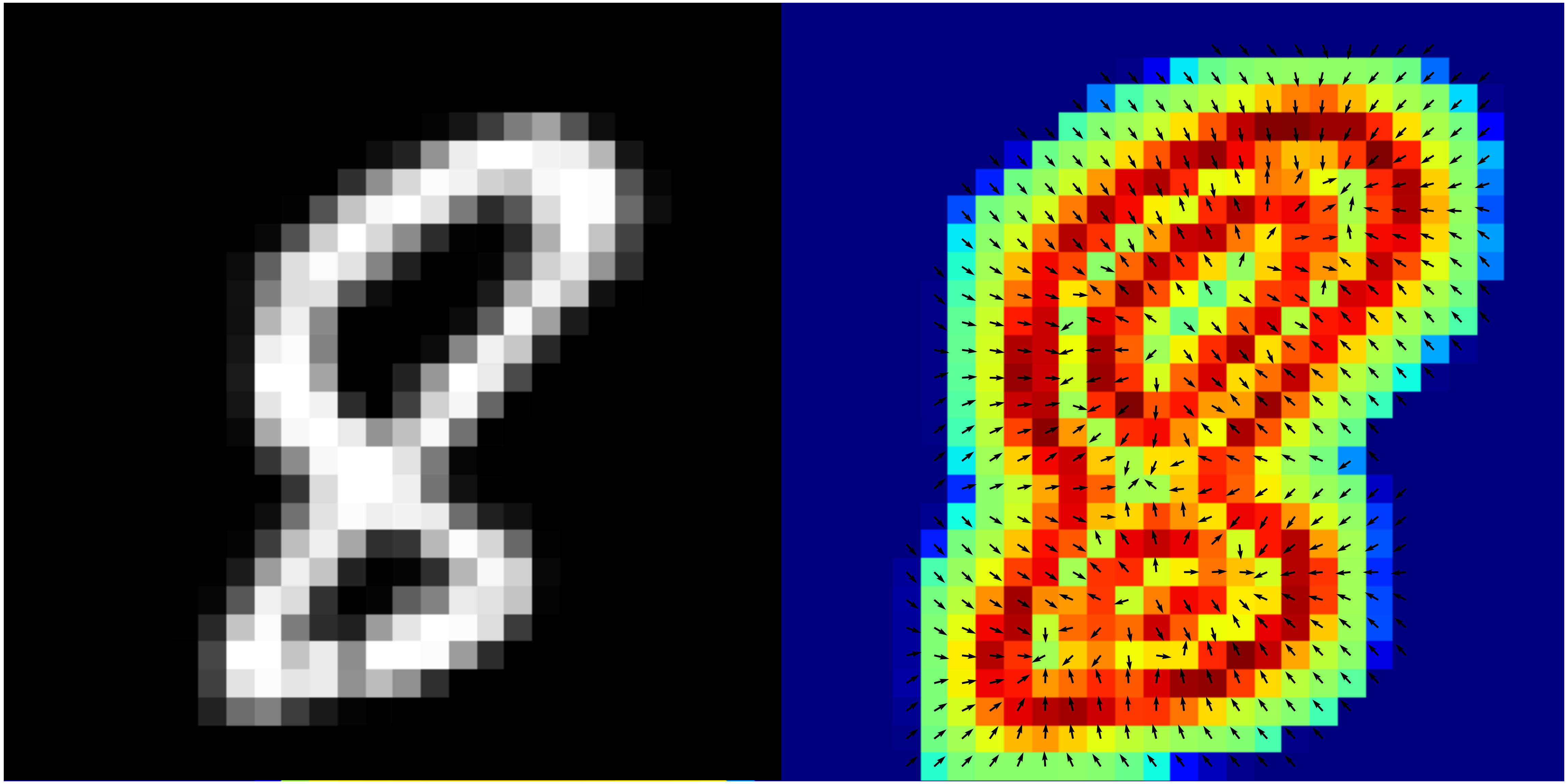
$$W_m \star I = \underbrace{(W_m^{\mathrm{Re}} \star I^{\mathrm{Re}} - W_m^{\mathrm{Im}} \star I^{\mathrm{Im}})}_{\text{Real}} + \iota \underbrace{(W_m^{\mathrm{Re}} \star I^{\mathrm{Im}} + W_m^{\mathrm{Im}} \star I^{\mathrm{Re}})}_{\text{Imaginary}}$$

CIRCULAR HARMONIC WEIGHTS



CIRCULAR HARMONICS

X-correlation w. circular harmonics returns orientation field



CIRCULAR HARMONIC ALGEBRA

So far we have a mapping $W\star : \mathbf{R}^{M\times N\times K} \to \mathbf{C}^{M\times N\times 1}$

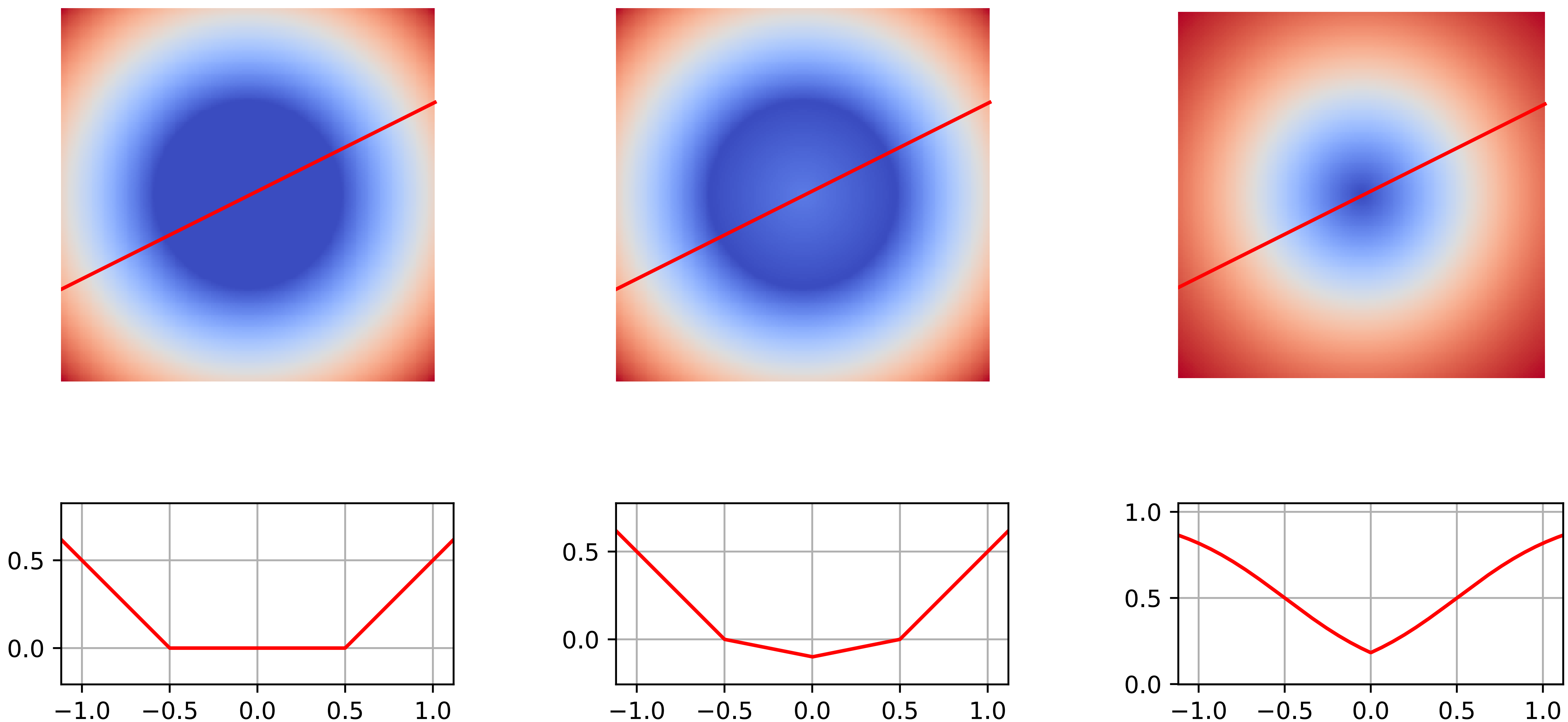
But what about:
1) nonlinearities?
2) subsequent layers?

$$W_{m_2} \star g([W_{m_1} \star \mathcal{T}_\theta I])$$

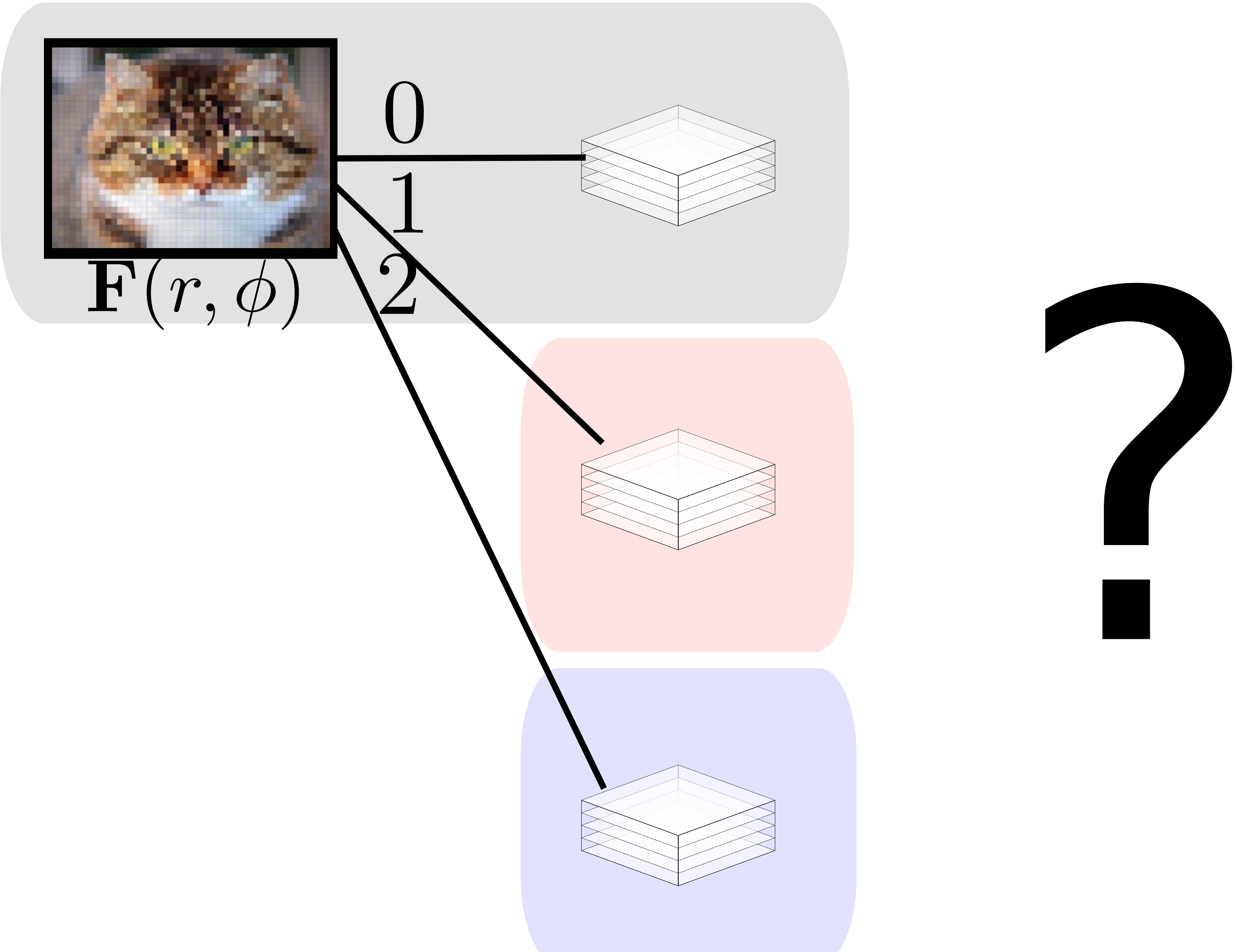
NONLINEARITIES

To maintain linear equivariance apply nonlinearities to invariant magnitude only $g(|F|e^{\iota\angle F}) = g(|F|)e^{\iota\angle F}$

$\mathrm{ReLU}(|F|+c)$ $\mathrm{LeakyReLU}(|F|+c)$ $\sigma(|F|+c)$



CIRCULAR HARMONIC ALGEBRA

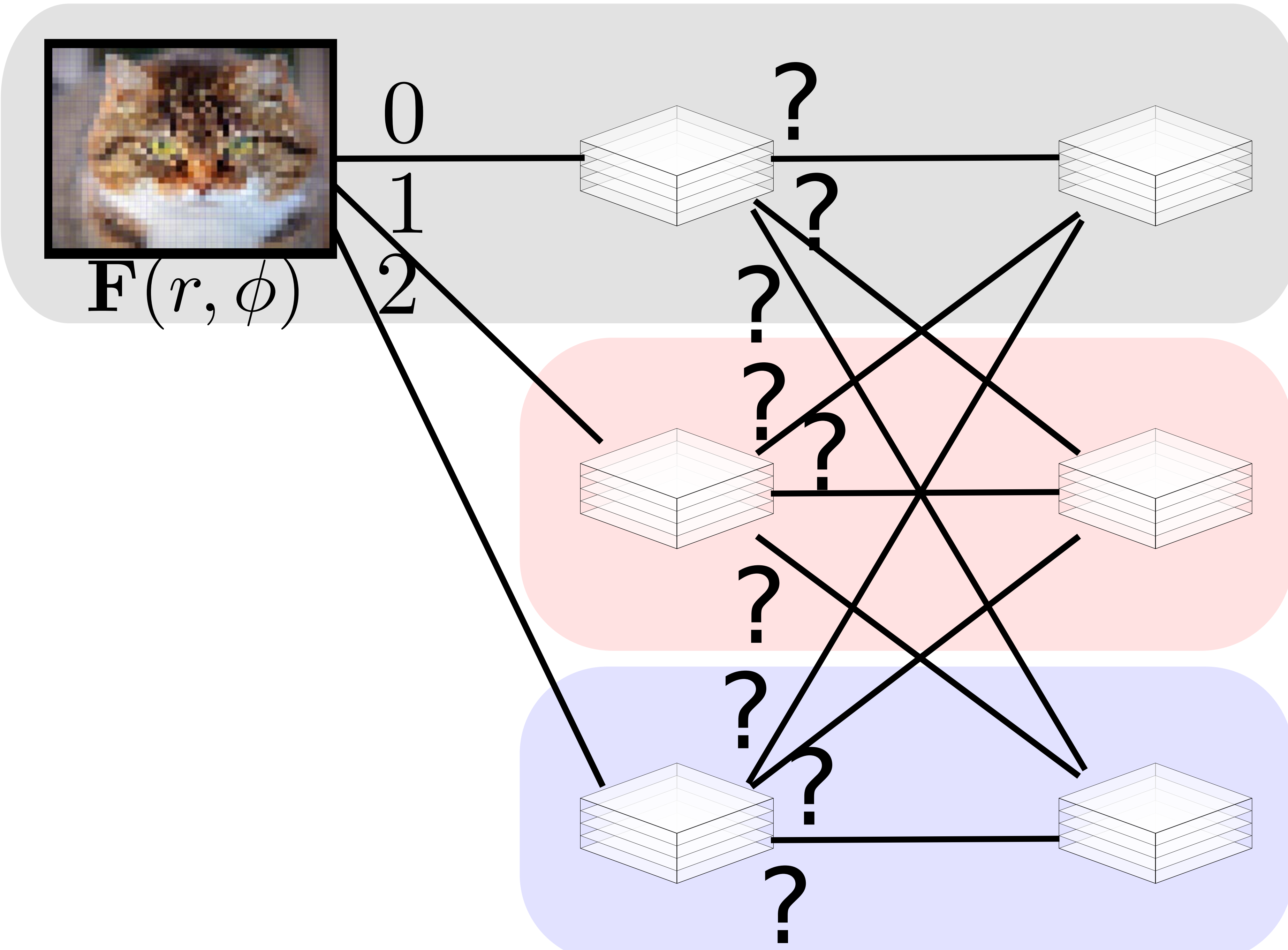


CIRCULAR HARMONIC ALGEBRA

Sum orders along chained cross-correlations

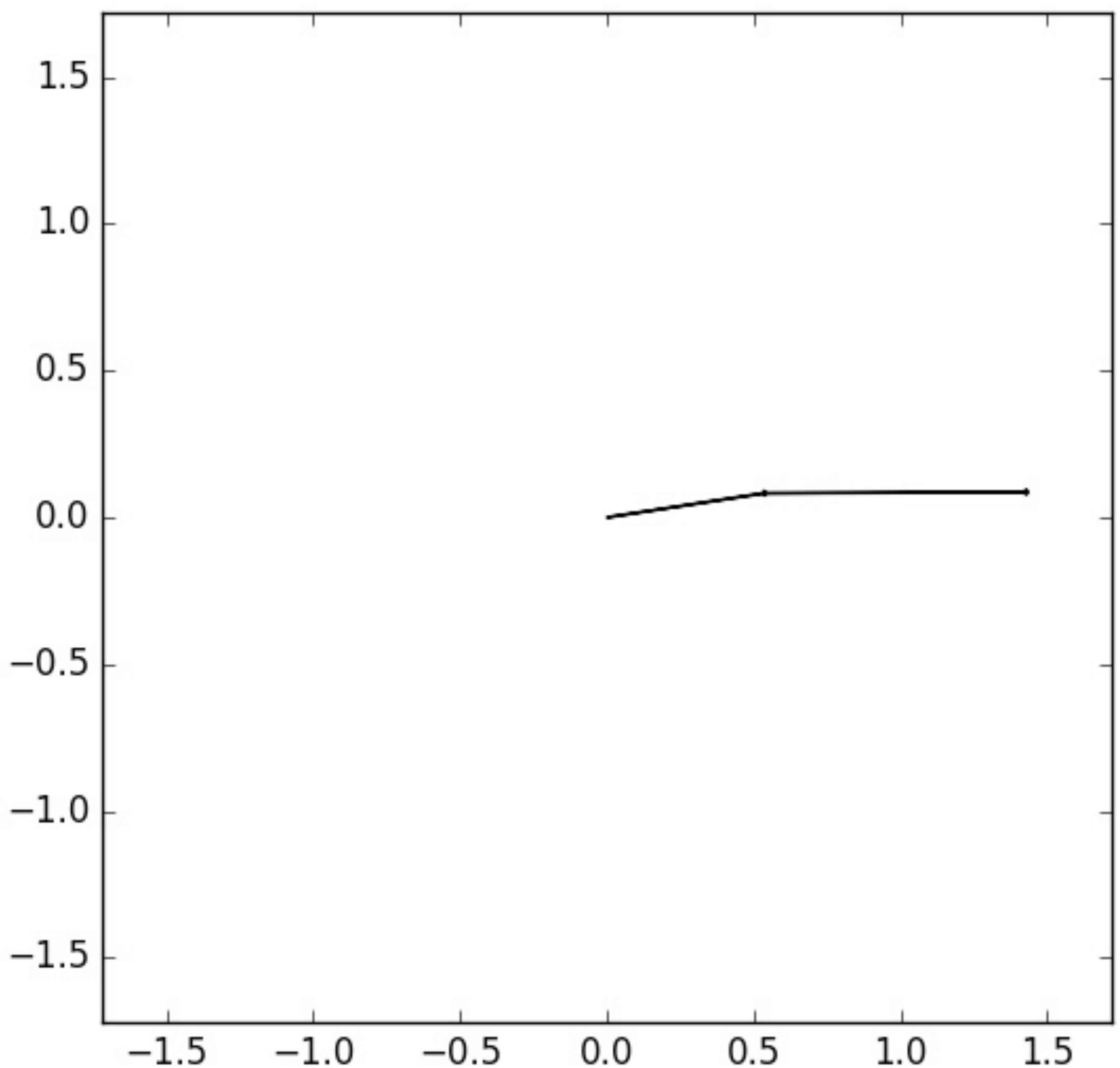
$$W_{m_2} \star g([W_{m_1} \star \mathcal{T}_\theta I]) = Y e^{\iota (m_1 + m_2)\theta}$$

CIRCULAR HARMONIC ALGEBRA



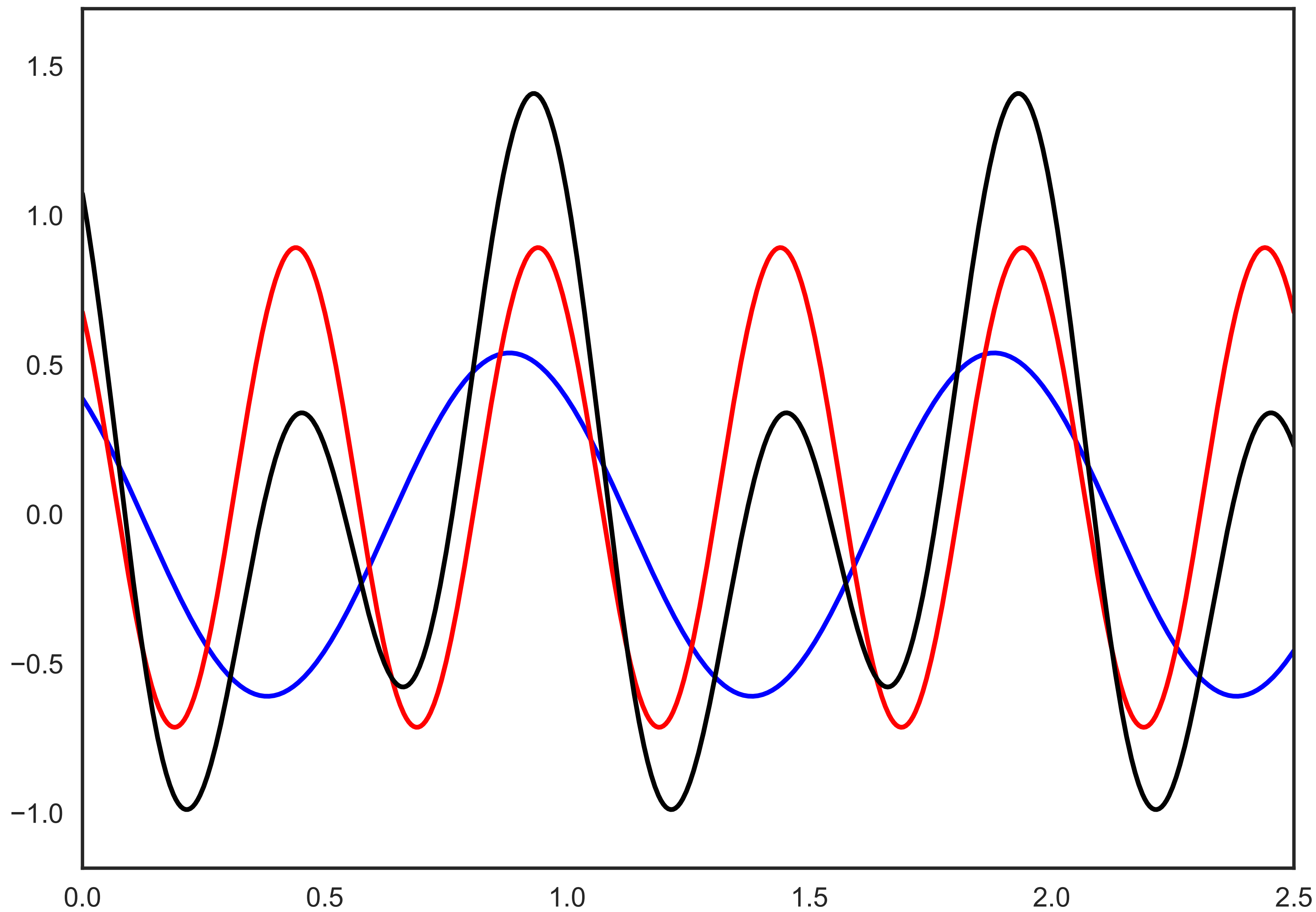
CIRCULAR HARMONIC ALGEBRA

At a layer, we sum over incoming feature maps, but need to pay special attention to rotation order



CIRCULAR HARMONIC ALGEBRA

At a layer, we sum over incoming feature maps, but need to pay special attention to rotation order



CIRCULAR HARMONIC ALGEBRA

The Equivariance Condition

Rotation order of filter

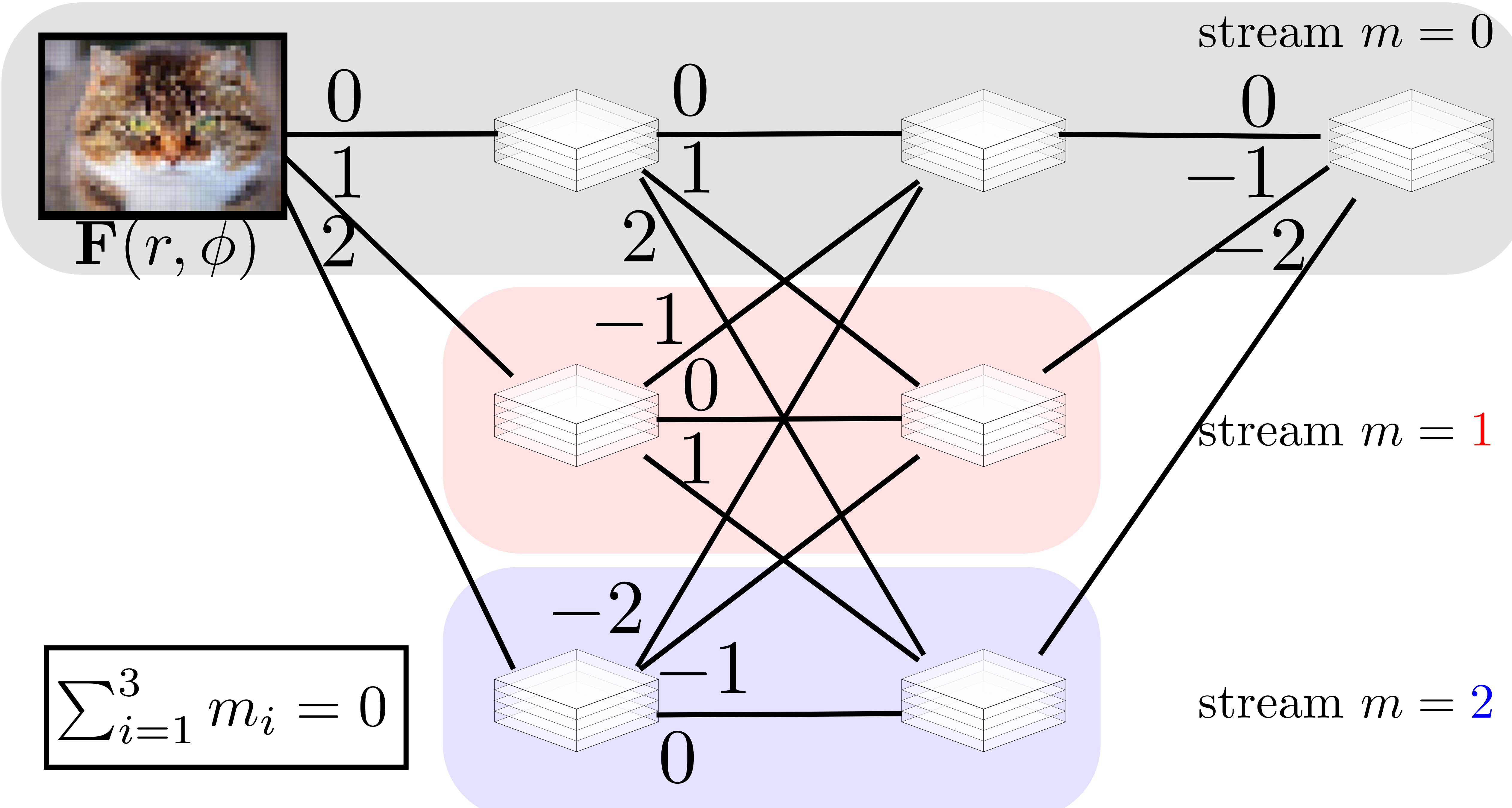
Constant at feature map i

$$\sum_{p \in \mathcal{P}_i} m_p = \kappa_i$$

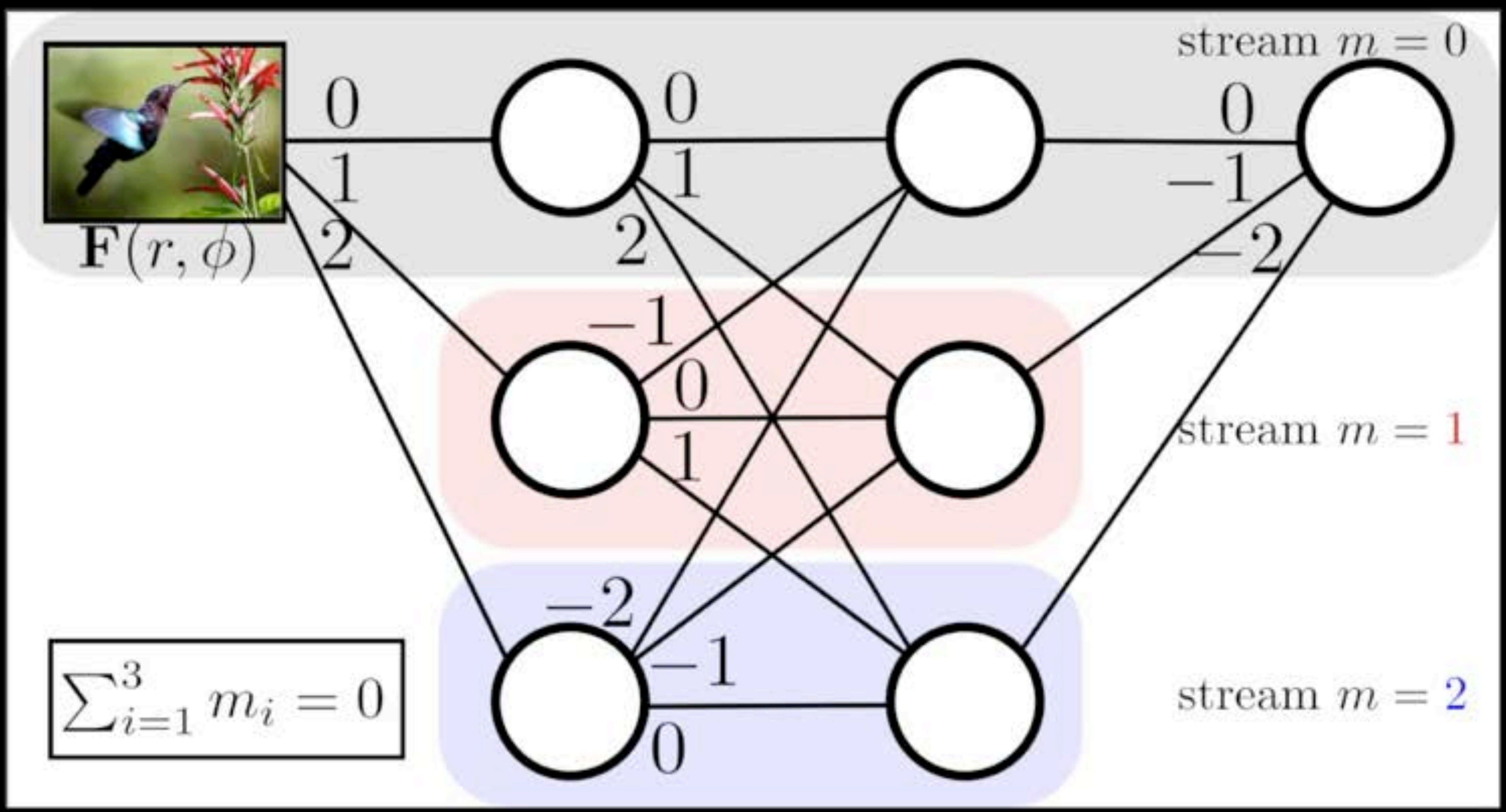
Path through i

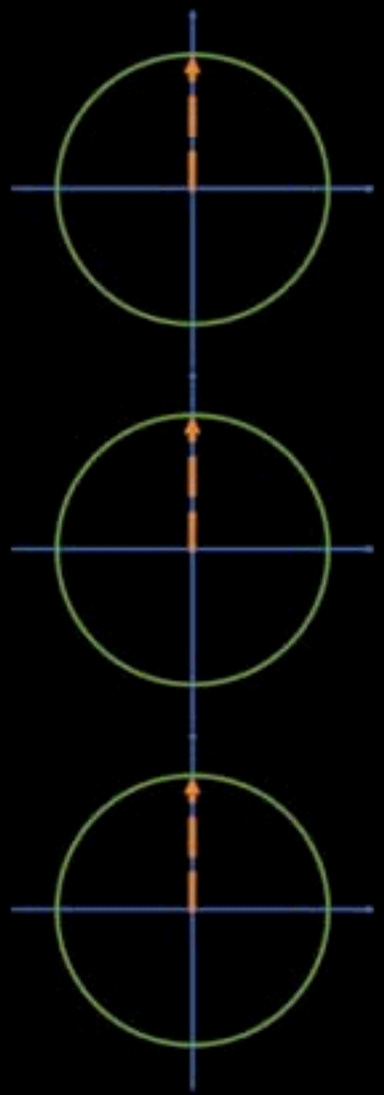
My conjecture: For Lie Groups, streams identified with basis vectors of Lie Algebra, and harmonic filters act as raising/lowering operators, jumping between basis vectors

CIRCULAR HARMONIC ALGEBRA

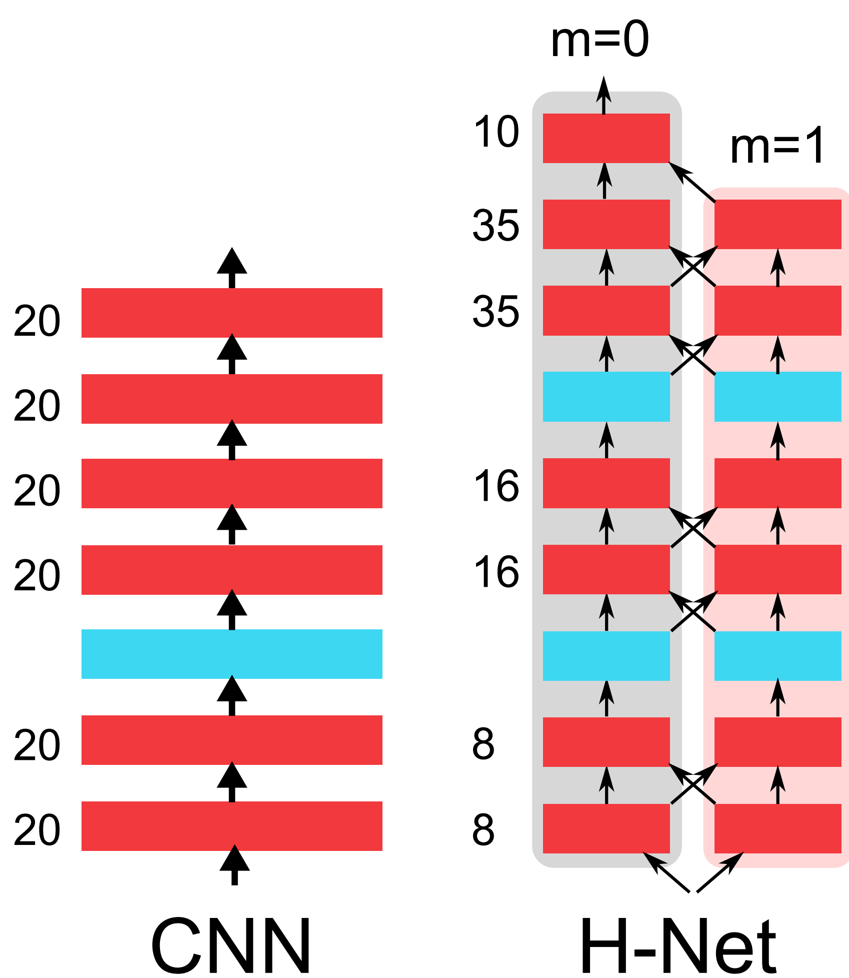


CIRCULAR HARMONIC ALGEBRA





SOME EXPERIMENTS



Brute force

Rotation + Scale

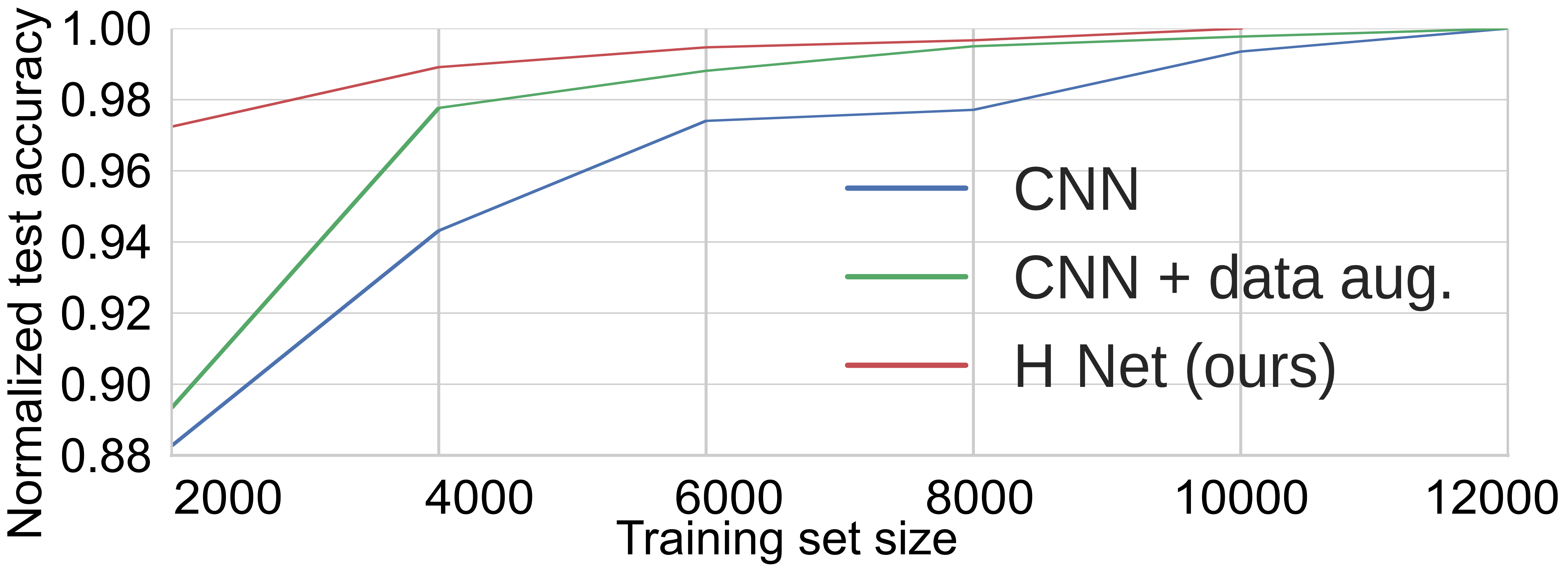
Sanity check: MNIST-*rot*

Method	Test error (%)	#params
SVM [1]	11.11	-
Transformation RBM [2]	4.2	-
Conv-RBM [3]	3.98	-
CNN [4]	5.03	22k
CNN + data aug* [4]	3.50	22k
P4CNN rotation pooling [4]	3.21	25k
P4CNN [4]	2.28	25k
Harmonic Networks (Ours)	1.69	33k

Recent results

ORN-8 [30]	1.37	~1M
TI-Pooling [13]	1.21	~1M
PTN-CNN-B++	0.89	~254k

SOME EXPERIMENTS



SOME EXPERIMENTS

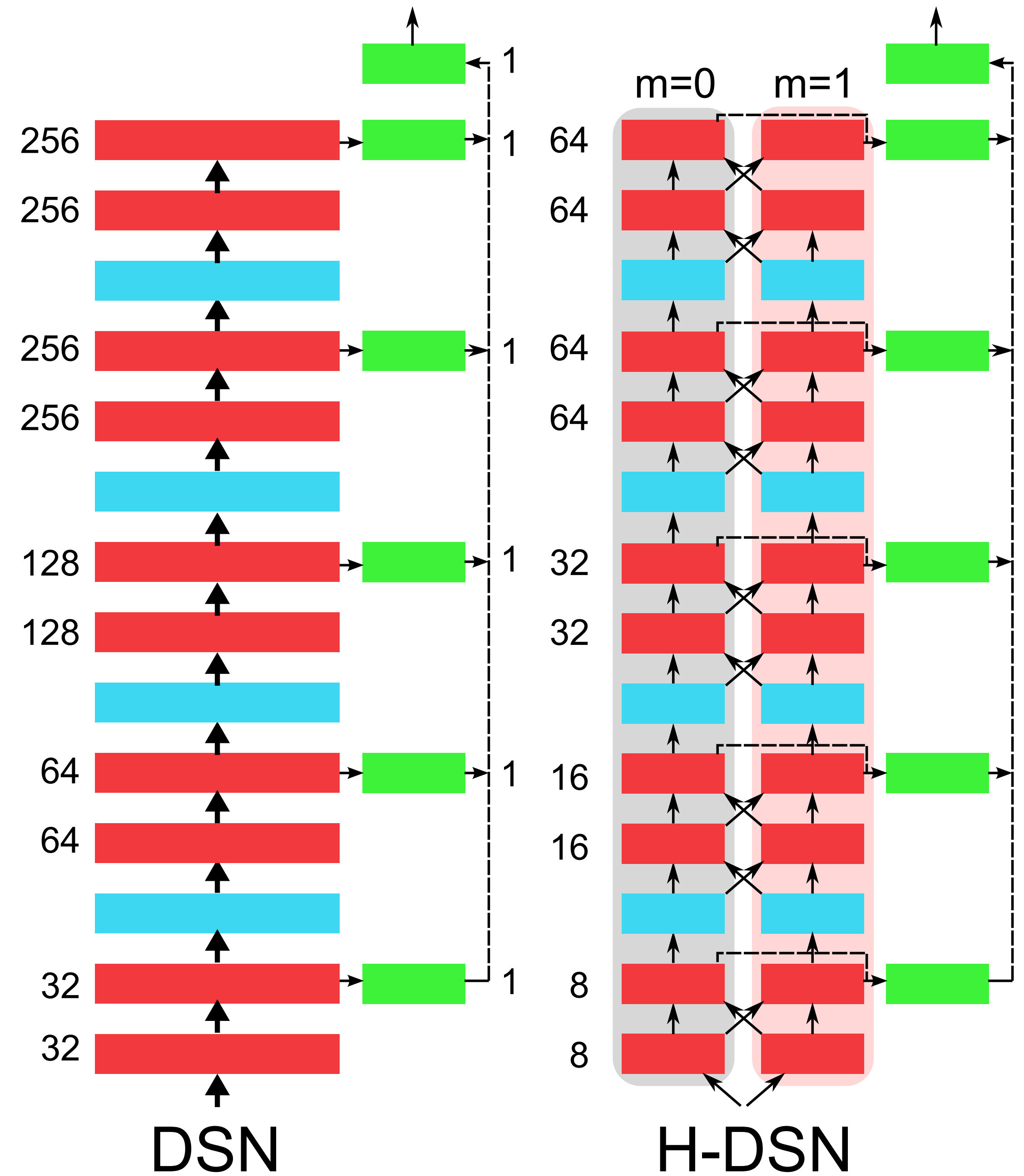
Boundary segmentation: BSD500

Feature magnitudes: w/o transfer learning

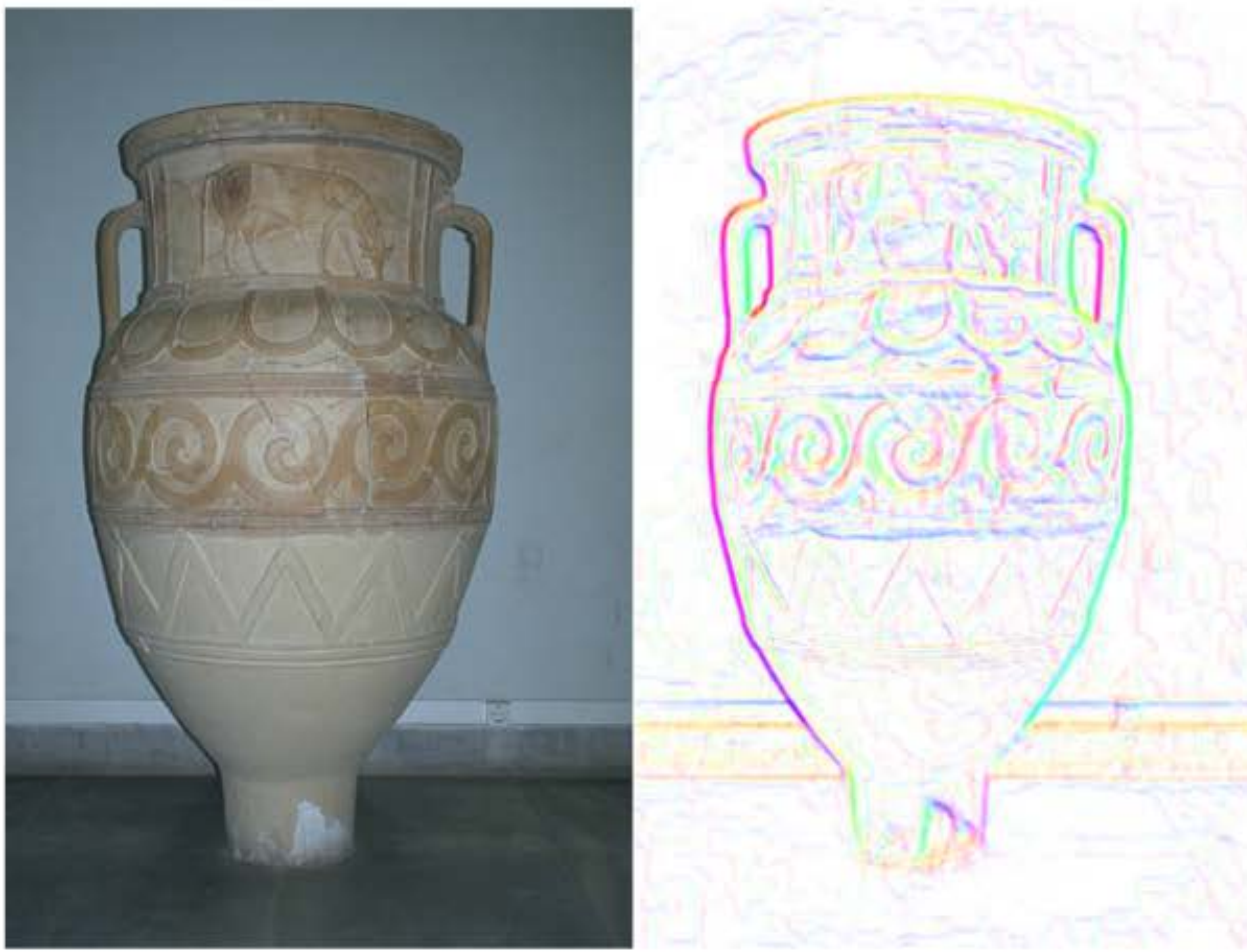
Method	ODS	OIS	#params
Holistically-nested edge detection (HED) [5]	0.640	0.650	2346k
HED (low # params) [5]	0.697	0.709	115k
Kivinen et al. [6]	0.702	0.715	-
Harmonic Networks (Ours)	**0.726**	**0.742**	116k

Recent results

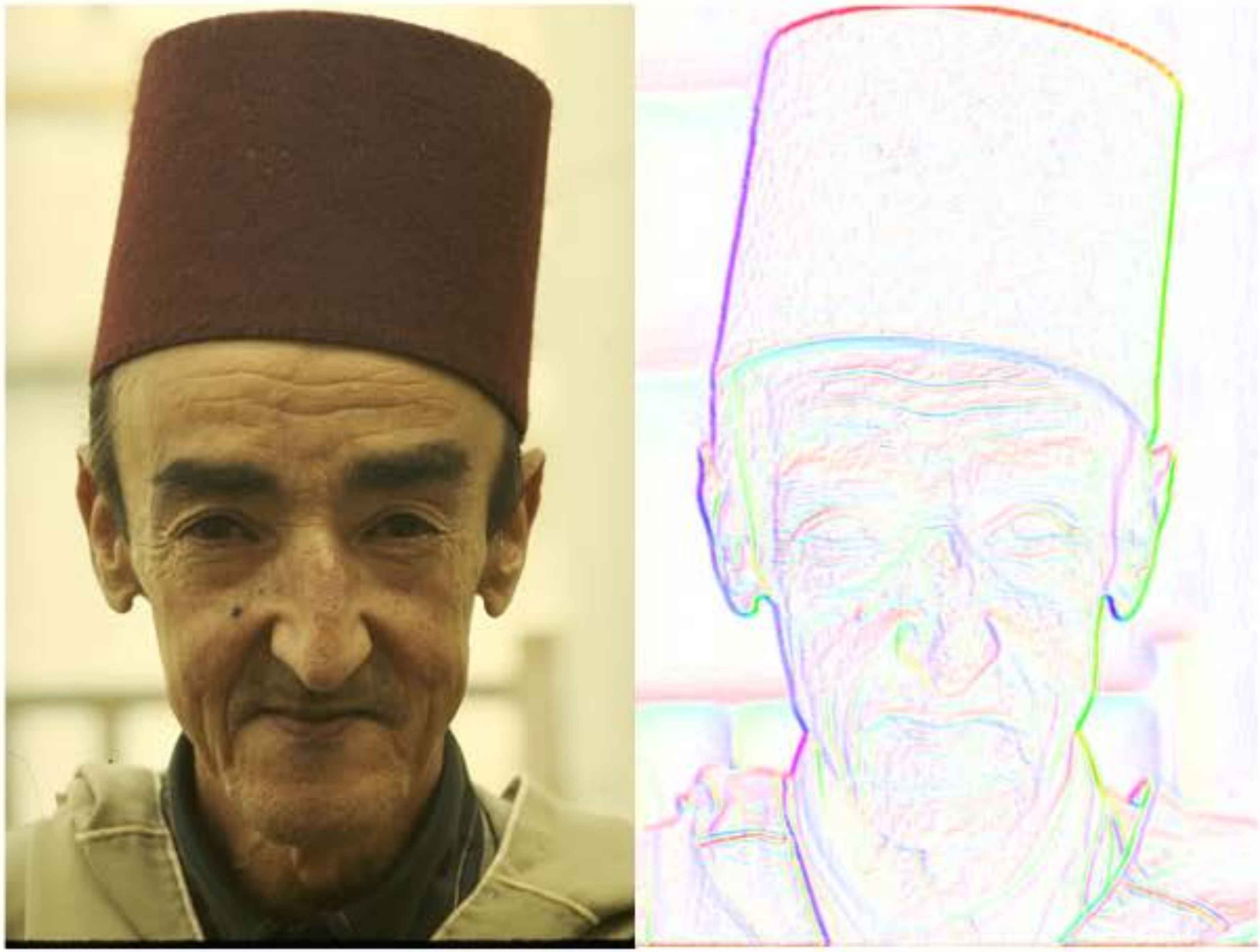
Dynamic Steerable ResNet **0.732 0.751**



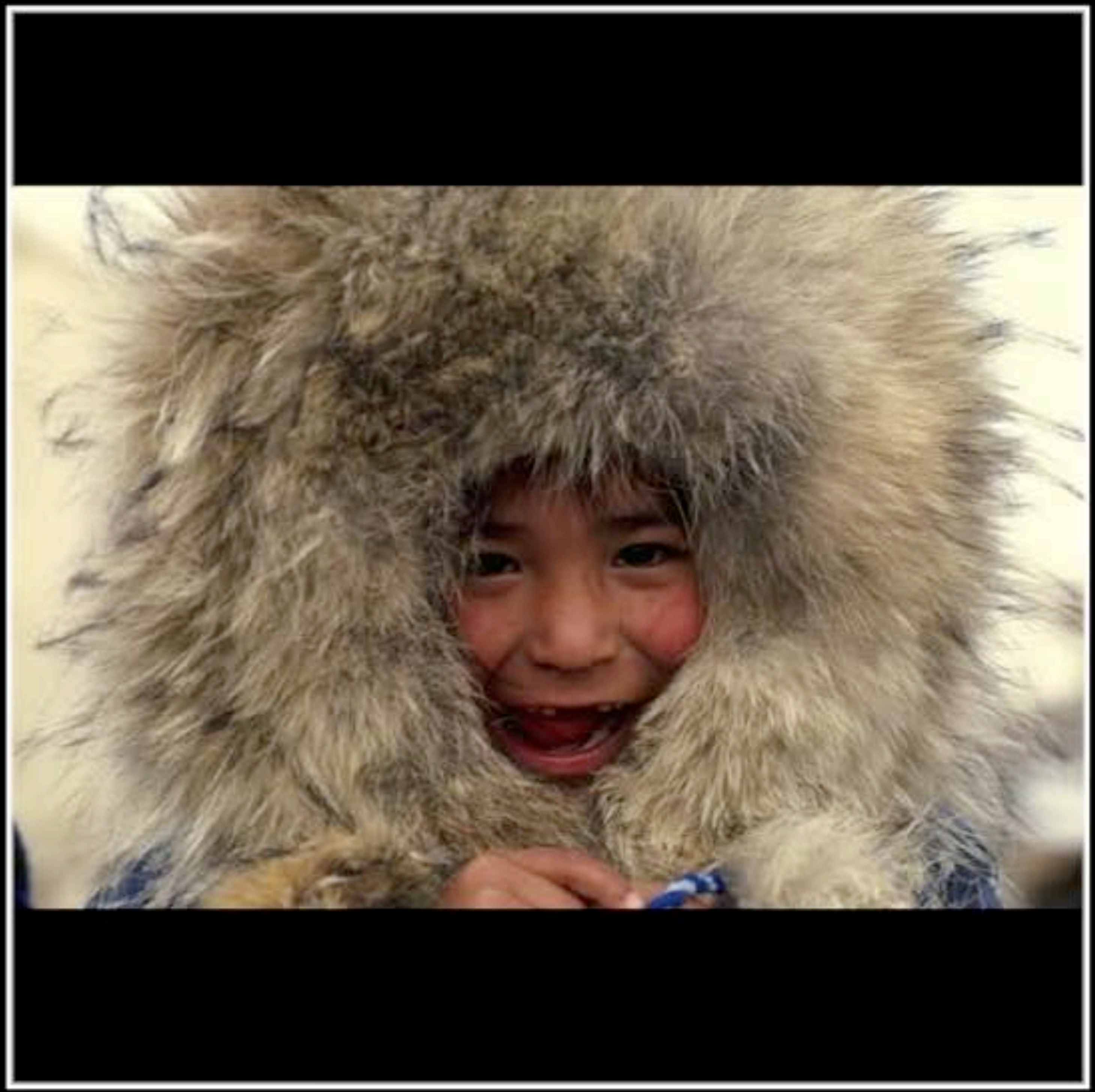
SOME EXPERIMENTS

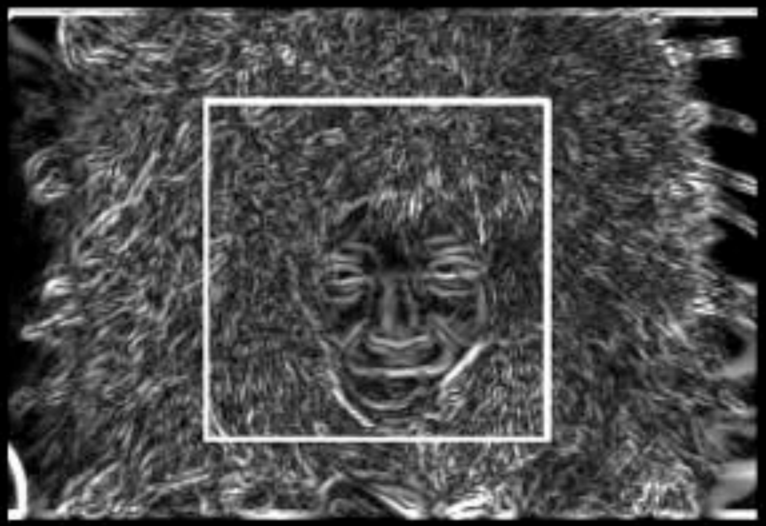


SOME EXPERIMENTS

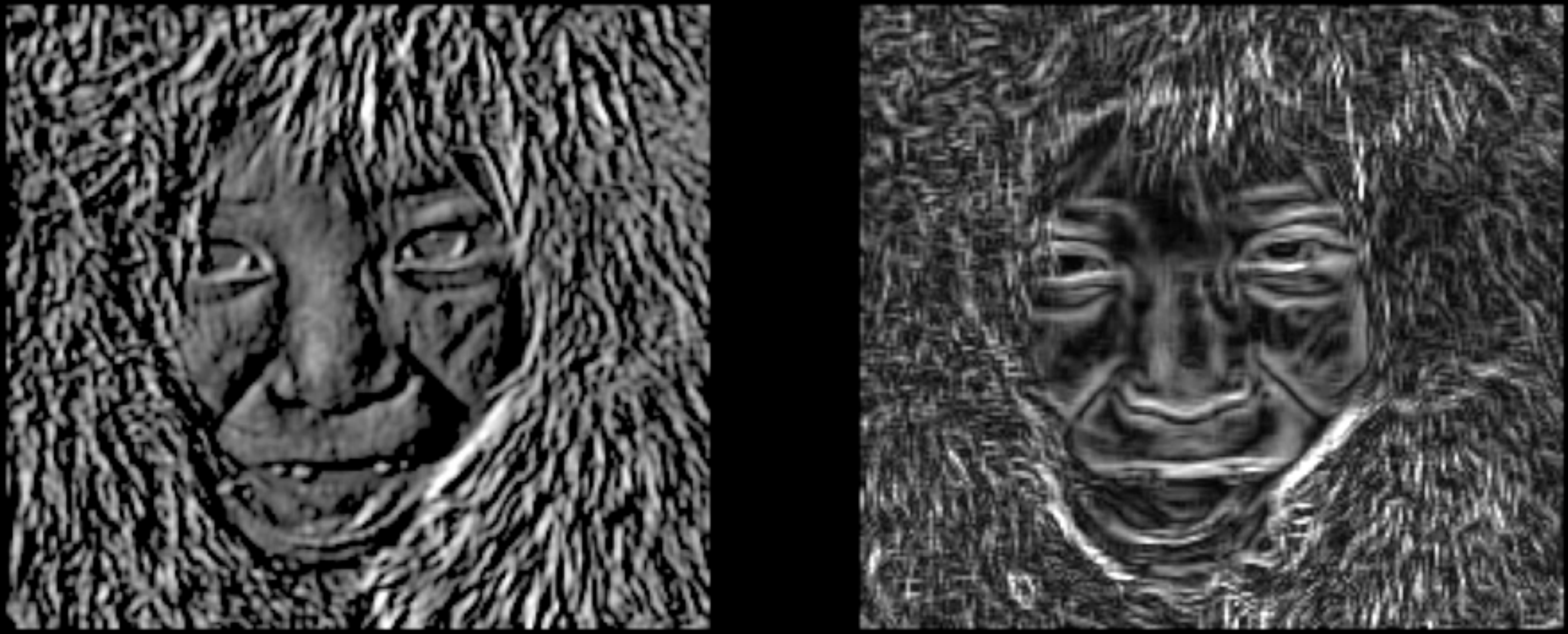


SOME EXPERIMENTS





SOME EXPERIMENTS



LIMITATIONS

- Filters too basic
- No stability guarantee for noise/high freq. deformations
- Do we really need higher order streams?
- What about other kinds of transformation?

Working on all of these

REFERENCES

Rotate filters

- Deep Roto-Translation Scattering for Object Classification *(Oyallon et al., 2014)*
- Group Equivariant Convolutional Networks *(Cohen et al., 2016)*
- Rotation Equivariant Vector Field Network *(Marcos et al., 2017)*
- Oriented Response Networks *(Zhou et al., 2017)*
- Steerable CNNs *(Cohen et al., 2017)*
- Roto-Translation Invariant Scattering *(Sifre et al., 2013)*

Rotate activations

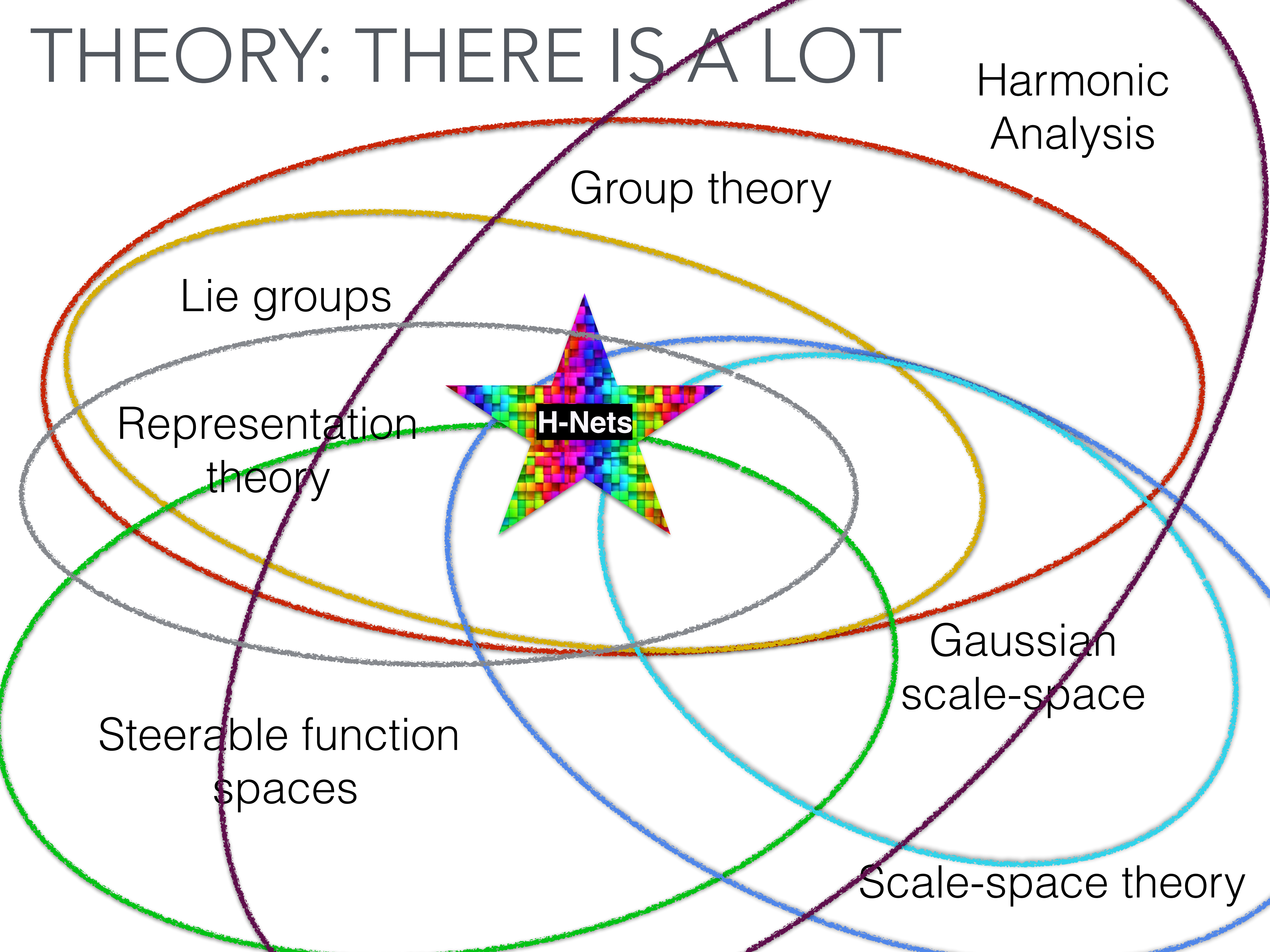
- Exploiting Cyclic Symmetry in Convolutional Neural Networks *(Dieleman et al., 2016)*

Rotate input

- Rotation-Invariant Neoperceptron *(Fasel et al., 2006)*
- TI-POOLING: Transformation-invariant Pooling for Feature Learning in Convolutional Neural Networks *(Laptev et al., 2016)*
- Warped Convolutions: Efficient Invariance to Spatial Transformations *(Henriques et al., 2017)*
- Polar Transformer Networks *(Esteves et al., 2017)*

PART III:
THEORY

THEORY: THERE IS A LOT



THE FOURIER SHIFT THEOREM ON LIE GROUPS

Theory is very involved, so this is just a brief outline

GROUP THEORY

A group $\mathfrak{G}$=(G, •) is a set G and a binary composition operator •, with

1. Closure

$$g \bullet h \in G$$

2. Associativity

$$(f \bullet g) \bullet h = f \bullet (g \bullet h)$$

3. Identity element

$$g \bullet e = e \bullet g = g$$

4. Unique inverses

$$g \bullet g^{-1} = g^{-1} \bullet g = e$$

GROUP THEORY: EXAMPLE

$(\mathbf{C} \setminus \{0\}, \cdot)$

1. Closure	$y \cdot z \in \mathbf{C} \setminus \{0\}$
2. Associativity	$(x \cdot y) \cdot z = x \cdot (y \cdot z)$
3. Identity element	$x \cdot 1 = 1 \cdot x = x$
4. Unique inverses	$x^{-1} = \frac{1}{\lvert x \rvert} e^{-\iota \angle x}$

2D ROTATIONS FORM A GROUP

These are some different ways to describe rotations

$$([0, 2\pi), +_{mod\,2\pi})$$

$$(\tilde{\mathbf{C}}, \cdot)$$

$$\tilde{\mathbf{C}} = \{z \in \mathbf{C} : |z| = 1\}$$

SO(2)

$$(\{\mathbf{A} \in \mathbf{M}^{2\times 2} : \mathbf{A}^\top \mathbf{A} = \mathbf{A}\mathbf{A}^\top = \mathbf{I}, |\mathbf{A}| = 1\}, \cdot)$$

There are an infinite number of *representations* to describe essentially the same thing

In fact this group has a manifold structure: it is a Lie Group

(Don’t worry if none of this makes sense—I only understand it every 2nd day)

PATCH IS FUNCTION ON SO(2)

FOURIER TRANSFORMS ON LIE GROUPS

Previously $\int f(\phi) e^{-\iota m \phi} \, \mathrm{d}\phi$

$$e^{-\iota m \phi} e^{-\iota m \theta} = e^{-\iota m (\phi + \theta)}$$

Now $\int f(\phi) \Psi_m(\phi) \, \mathrm{d}\phi$

Homomorphism property

$$\Psi_m(\theta) \Psi_m(\phi) = \Psi_m(\theta \bullet \phi)$$

(We lose commutativity)

$$\int_G \mathcal{T}_\theta[f](\phi) \Psi_m(\phi) \, \mathrm{d}\phi = \int_G f(\theta^{-1} \phi) \Psi_m(\phi) \, \mathrm{d}\phi$$

Need left-invariant Haar measure

$$= \int_{\theta^{-1}\tilde{G}} f(\tilde{\phi}) \Psi_m(\theta \tilde{\phi}) \, \mathrm{d}(\theta \tilde{\phi})$$

Extension of Fourier Shift Theorem

$$= \Psi_m(\theta) \int_{\tilde{G}} f(\tilde{\phi}) \Psi_m(\tilde{\phi}) \, \mathrm{d}\tilde{\phi}$$

Equivariant

Invariant

INVARIANTS ON LIE GROUPS

Can show that the irreducible representations are unitary so

$$\mathcal{F}_m\{f\} := \int_G f(\phi)\Psi_m(\phi)\,\mathrm{d}\phi \qquad \|H\|^2 = \langle H|H\rangle = \mathrm{trace}(\tilde{H}^\top H)$$

$$\begin{aligned}
\|\mathcal{F}_m\{\mathcal{T}_\theta f\}\|^2 &= \|\Psi_m(\theta)\mathcal{F}_m\{f\}\|^2 \\
&= \langle \Psi_m(\theta)\mathcal{F}_m\{f\}|\Psi_m(\theta)\mathcal{F}_m\{f\}\rangle \\
&= \mathrm{trace}(\widetilde{\mathcal{F}_m\{f\}}^\top \underbrace{\widetilde{\Psi_m(\theta)}^\top \Psi_m(\theta)}_{I}\,\mathcal{F}_m\{f\}) \\
&= \|\mathcal{F}_m\{f\}\|^2
\end{aligned}$$

Invariant on group

In such a way we can proceed as before to define a harmonic network over a matrix Lie group

WORK IN PROGRESS

Open questions:

- How to jump between streams?
- Transformations such as scalings are *non-compact*, can we approximate them well on an interval?
- Not all transformations form groups. Harmonics are eigenfunctions of the Laplace-Beltrami operator.
    - Could we define equivariance on general manifolds or even graphs?
    - What about eigenfunctions of other operators? c.f. Sturm-Liouville theory
- There are other equivariant transforms such as the Gaussian convolution (affine equivariant). This has a semi-group structure, how does this fit in?

THANK YOU