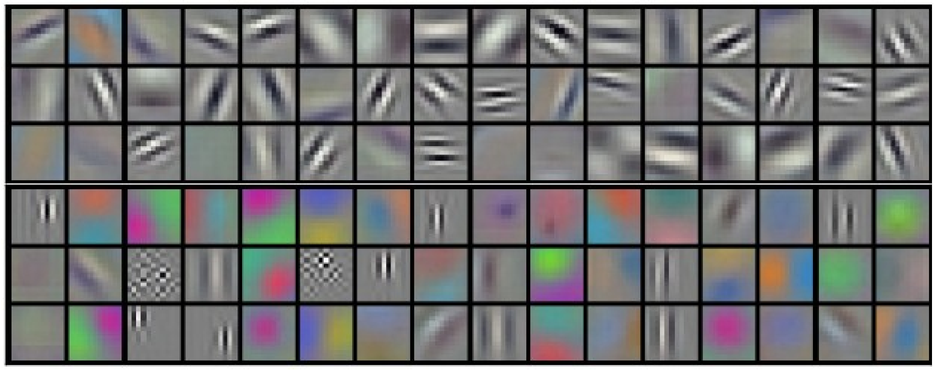
TTIC 31230, Fundamentals of Deep Learning

David McAllester, Winter 2018

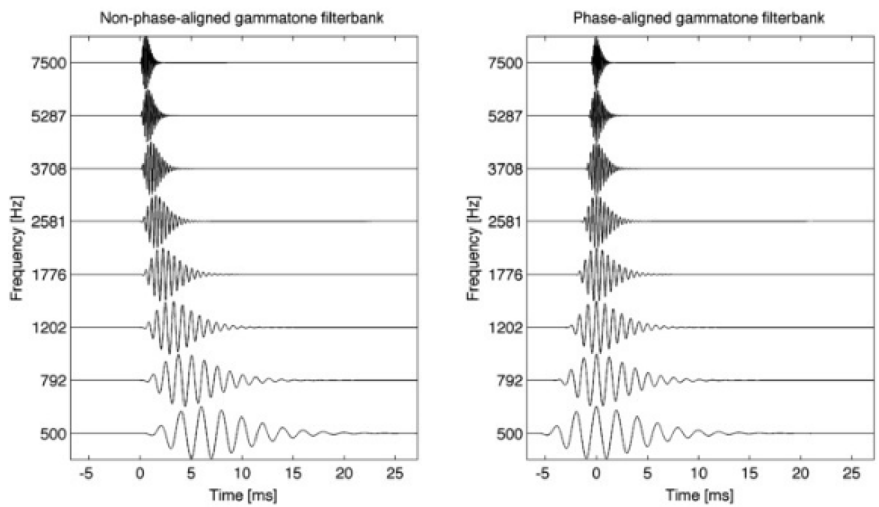
Invariant Theory

Invariant Theory

Why Are Early Filters Wavelets?



Krizhevsky



MathWorks

Invariance

Consider the distribution of “natural” 360 degree images.

It should be true that the probability of a given 360 degree image is equal to the probability of any rotation of that image.

We say that the probability distribution is invariant to rotation.

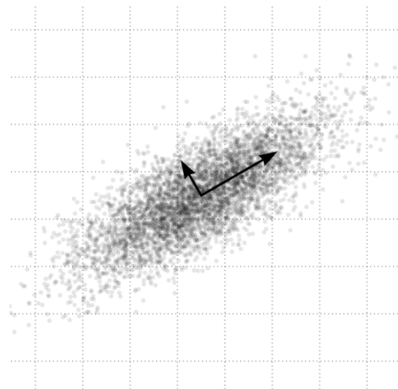
Invariances

Translation invariance (in both space and time)

Scale invariance (in both space and time)

Rotation invariance (spatial rotations).

PCA and Invariance



The principal components in PCA are the eigenvectors of the covariance matrix.

The principal components of the covariance matrix of a translation invariant distribution are the Fourier basis functions (sine and cosine).

PCA and Invariance

The eigenvalues of the covariance matrix are given by the power spectrum of the signal distribution.

This is the Einstein-Wiener-Khinchin theorem (proved by Wiener, and independently by Khinchin, in the early 1930s, but — as only recently recognized – stated by Einstein in 1914). From "Signals and Systems" by Oppenheim and Verghese

This explains projection onto complex exponentials as a first step in signal processing and signal compression (e.g., JPEG).

More Formally

Let ρ be a probability density over vectors in $\mathbb{R}^n$.

We say ρ is rotation-stationary if

- $\mathrm{E}\left[x_i\right] = \mathrm{E}\left[x_j\right]$ for all i, j.
- $\mathrm{E}\left[x_i x_j\right] = f(i - j \bmod n)$

Rotation stationarity is a simplification of the more widely used notion of translation stationarity (or just stationarity).

More Formally

The covariance matrix is given by

$$\Sigma_{i,j} = \mathrm{E}\left[x_i x_j - \mathrm{E}\left[x_i\right]\mathrm{E}\left[x_j\right]\right] = g(i - j \bmod n)$$

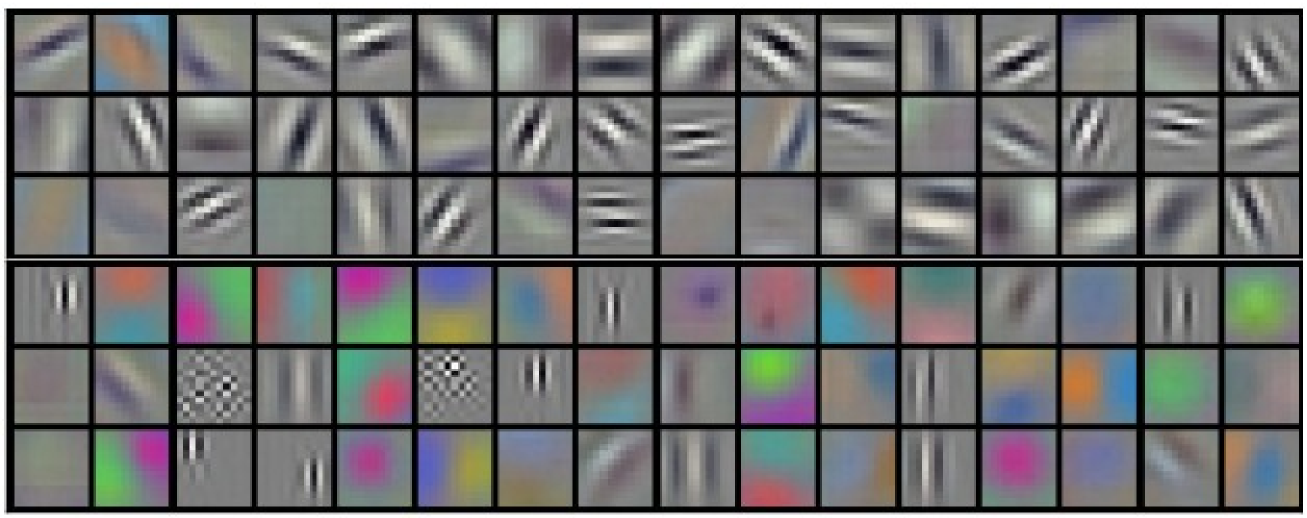
A matrix satisfying $\Sigma_{i,j} = g(i-j \bmod n)$ is called **circulant**.

The eigenvectors of a circulant matrix form a discrete Fourier basis.

Wavelets

In practice we want the compressed representation to be local to also satisfy scale invariance. This leads to **wavelets**.

To my knowledge scale invariance is not currently built into deep vision architectures.



Invariance and Data Augmentation

CNNs build translation invariance into the architecture.

Another approach to invariance is to apply invariant transformations to the training data.

For example we can apply translations, scalings, rotations, reflections to generate more labeled images in MNIST or Imagenet to get a much larger training set.

END