Computing Perceptual Invariants

Sophia Sanborn
VS265 Guest Lecture
Fall 2020
Outline

1. The problem of invariance
2. The mathematics of invariance
3. Invariance in visual cortex
4. Invariance in convolutional neural networks
The Problem of Invariance
The Problem of Invariance
The Problem of Invariance
THE MONIST

Henri Poincaré

ON THE FOUNDATIONS OF GEOMETRY.

ALTHOUGH I have already had occasion to set forth my views on the foundations of geometry, it will not, perhaps, be unprofitable to revert to the question with new and ampler developments, and seek to clear up certain points which the reader may have found obscure. It is with reference to the definition of the point and the determination of the number of dimensions that new light appears to me most needed; but I deem it opportune, nevertheless, to take up the question from the beginning.

SENSIBLE SPACE.

Our sensations cannot give us the notion of space. That notion is built up by the mind from elements which pre-exist in it, and external experience is simply the occasion for its exercising this power, or at most a means of determining the best mode of exercising it.
The Mathematics of Invariance
The Mathematics of Invariance
The Mathematics of Invariance

Invariance

\[ f(\phi(x)) = f(x) \]
The Mathematics of Invariance

Invariance

\[ f(\phi(x)) = f(x) \]
The Mathematics of Invariance

Invariance

\[ f(\phi(x)) = f(x) \]

Equivariance

\[ f(\phi(x)) = \phi'(f(x)) \]

\[ F \]
The Mathematics of Invariance

**Invariance**

\[ f(\phi(x)) = f(x) \]

**Equivariance**

\[ f(\phi(x)) = \phi'(f(x)) \]
The Mathematics of Invariance

Invariance
\[ f(\phi(x)) = f(x) \]

Equivariance
\[ f(\phi(x)) = \phi'(f(x)) \]
The Mathematics of Invariance

\[ f_n \]
The Mathematics of Invariance

\[ f_n = \]
The Mathematics of Invariance

\[ f_n = \sum_{n=0}^{N-1} \hat{f}_\omega e^{\frac{i2\pi\omega n}{N}} \]
The Mathematics of Invariance

\[ f_n = \sum_{n=0}^{N-1} \hat{f}_\omega e^{i2\pi n \omega n/N} \]

\[ e^{i\theta} = \cos \theta + i \sin \theta \]

Euler’s Formula
The Mathematics of Invariance

\[ f_n = \sum_{n=0}^{N-1} \hat{f}_\omega e^{i2\pi n \omega \frac{n}{N}} \]
The Mathematics of Invariance

\[ f_n = \sum_{n=0}^{N-1} \hat{f}_\omega e^{\frac{i2\pi n\omega}{N}} \]
The Mathematics of Invariance

\[ f_n = \sum_{n=0}^{N-1} \hat{f}_\omega e^{\frac{i2\pi on}{N}} \]

\[ \hat{f}_\omega = \sum_{n=0}^{N-1} f_n e^{-\frac{i2\pi on}{N}} \]

The Discrete Fourier Transform
The Mathematics of Invariance

The Discrete Fourier Transform

\[ f_n = \sum_{n=0}^{N-1} \hat{f}_\omega e^{\frac{i2\pi on}{N}} \]

\[ \hat{f}_\omega = \sum_{n=0}^{N-1} f_n e^{-\frac{i2\pi on}{N}} \]
The Mathematics of Invariance
The Mathematics of Invariance
The Mathematics of Invariance

Original Signal

Shifted by $\pi/2$
The Mathematics of Invariance

Top Fourier Coefficients
The Mathematics of Invariance

Original Fourier Coefficients

Fourier Coefficients of Shifted Signal
Power Spectrum

\[ P_{\omega} = |\hat{f}_{\omega}|^2 \]
The Mathematics of Invariance
The Mathematics of Invariance
The Mathematics of Invariance
Invariance in Visual Cortex

Simple Cells

Complex Cells

Stimulus: on off
Invariance in Visual Cortex
Invariance in Visual Cortex

Quadrature Pair of Gabors
Invariance in Visual Cortex

Quadrature Pair of Gabors

The Energy Model
The Mathematics of Invariance
The Mathematics of Invariance

Bispectrum

\[ \beta(f)_{j,k} = \hat{f}_j^* \hat{f}_k^* \hat{f}_{j+k} \]
The Mathematics of Invariance

**Bispectrum**

\[ \beta(f)_{j,k} = \hat{f}_j \hat{f}_k \hat{f}_{j+k} \]

\[ \beta(f)_{j,k} = \left( \sum_{t=0}^{\tau} e^{i2\pi j t / \tau f_t} \right) \left( \sum_{t=0}^{\tau} e^{i2\pi k t / \tau f_t} \right) \left( \sum_{t=0}^{\tau} e^{-i2\pi (j+k) t / \tau f_t} \right) \]
The Mathematics of Invariance
Invariance in Computer Vision
Convolution

Pooling

Invariance in Computer Vision
Invariance in Computer Vision

Convolution

Pooling

Equivariance to translation within the global image plane
Invariance in Computer Vision

Convolution

Equivariance to translation within the global image plane
Invariance in Computer Vision

**Convolution**

Equivariance
to translation within the global image plane

**Pooling**

Invariance
to slight perturbations within a local region
Invariance in Computer Vision

Translation

Rotation
Invariance in Computer Vision

Max Pooling
Invariance in Computer Vision

Data Augmentation
Invariance in Computer Vision
Invariance in Computer Vision
Hierarchical Vision Models

All neurons in the previous layer with at least 30% of the max weight magnitude are shown, both **positive (excitation)** and **negative (inhibition)**. Click on a neuron to see its forwards and backwards weights.
Hierarchical Vision Models

- **Color Contrast** 16%
- **Complex Gabor** 14%
- **Low Frequency** 27%
<table>
<thead>
<tr>
<th>Feature</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angles</td>
<td>3%</td>
</tr>
<tr>
<td>Brightness Gradient</td>
<td>6%</td>
</tr>
<tr>
<td>BW vs Color</td>
<td>4%</td>
</tr>
<tr>
<td>Repeating patterns</td>
<td>5%</td>
</tr>
<tr>
<td>High-Low Frequency</td>
<td>6%</td>
</tr>
<tr>
<td>Curves</td>
<td>4%</td>
</tr>
</tbody>
</table>
Hierarchical Vision Models
InceptionV1 has a **left-oriented** pathway detecting dogs facing left...

...and a symmetric **right-oriented** pathway detecting dogs facing right. At each step, the two pathways inhibit each other and excite the next stage.

Union over left and right cases.

Orientation-Invariant Head (4b)

Orientation-Invariant Head+Neck (4c)
Too Invariant?

<table>
<thead>
<tr>
<th>king penguin</th>
<th>starfish</th>
<th>baseball</th>
<th>electric guitar</th>
</tr>
</thead>
<tbody>
<tr>
<td>freight car</td>
<td>remote control</td>
<td>peacock</td>
<td>African grey</td>
</tr>
</tbody>
</table>
Not Invariant
Enough?

Original image
Temple (97%)

Perturbations

Adversarial example
Ostrich (98%)