Learning Spatial and Transformational Invariants for Visual Representation

Charles Cadieu
work with Bruno Olshausen

CIFAR Summer School
Toronto
Visual Representation

Visual Cortex
(macaque monkey)

Charles Cadieu, cadieu@berkeley.edu
The Visual System Infers the Causes of Images

Transformational Invariants

“Causes”

Spatial Invariants
“Galloping” - Transformational Invariant

“Leopard” - Spatial Invariant
Structure within an Image Patch

- How do we uncover the causes of this complicated data?
\[ I(x,t) = \sum_i A_i(x)u_i(t) + \nu(x,t) \]
I(x,t) = φ_i(t) A_i^R(x) + \sin(φ_i(t)) A_i^I(x) + \nu(x,t)

Quickly Changing
 Sparse + Temporally Stable
First Layer
Probabilistic Model

\[ P(I, a, \phi) \propto e^{-E_1} \]

\[ E_1 = \sum_t \sum_x \frac{1}{\sigma^2} \left[ I(x,t) - \sum_i \Re\{z_i^*(t) A_i(x)\} \right]^2 + \lambda_{Sp} \sum_{i,t} a_i(t) + \lambda_{Sl} \sum_{i,t} (a_i(t) - a_i(t-1))^2 \]

Reconstruction Error
Sparse
Slow

Adapt to Natural Movies

Charles Cadieu, cadieu@berkeley.edu
Learned Basis Functions

\[ I \]

\[ A_{191}^R \]

\[ A_{191}^I \]

\[ u^R \]

\[ u^I \]

\[ a \]

\[ \phi \]
First Layer Basis Functions
Motions Produce Patterns in Phase

\[ \phi_i(t) \]
Model the changes in phase with a sparse, latent variable model

\[ \phi_i(t) - \phi_i(t-1) = \sum_k D_{ik} w_k(t) + \delta_i(t) \]
Second Layer Probabilistic Model:
Transformational Invariants

\[ P(I, a, \phi, w) \propto e^{-E_1 - E_2} \]

Error in Phase Dynamics

\[ E_2 = -\sum_t \sum_{i \in \{ a_i(t) > 0 \}} \kappa \cos(\dot{\phi}_i - [Dw(t)]_i) + \beta_{Sp} \sum_{k,t} |w_k(t)| + \beta_{Sl} \sum_{k,t} (w_k(t) - w_k(t-1))^2 \]

Adapt to Natural Movies

Charles Cadieu, cadieu@berkeley.edu
Visualizing Learned Weights

$D_{10}$

Spatial Position

Spatial Frequency

Charles Cadieu, cadieu@berkeley.edu
Transformation Components in Image Space
Transformation Components in Frequency Space
Learned Transformation Component

\[ D_{10} \]

\[ I \]

\[ u^R \]

\[ u^I \]

[Graph showing learned transformation components with nodes and arrows, including symbols for space and freq.]
Learned Transformational Invariant

\[ u \xrightarrow{a} \phi \xrightarrow{D_{10}} \text{freq.} \]

Charles Cadieu, cadieu@berkeley.edu

Wednesday, August 11, 2010
Learned Transformational Invariant

\[ \mathcal{W} \xrightarrow{D_{21}} \text{space freq.} \]

\[ \alpha \xrightarrow{} \phi \]

\[ u^R \xrightarrow{} u^I \]

Charles Cadieu, cadieu@berkeley.edu

Wednesday, August 11, 2010
Learned Transformational Invariant

$u$ $\xrightarrow{D_{24}}$ $\phi$ $\xrightarrow{\alpha}$ $u^R$, $u^I$

Charles Cadieu, cadieu@berkeley.edu
Second Layer Probabilistic Model: Spatial Invariants

\[ P(I, a, \phi, v) \propto e^{-E_1 - E_2} \]

Error in log(amplitude)

\[ E_2 = -\sum_i \frac{1}{2\sigma^2 M} (\log(a_i) - [Bv]_i)^2 + \beta_{Sp} \sum_k |v_k| \]

Sparse

Adapt to Natural Images

Charles Cadieu, cadieu@berkeley.edu
Learned Spatial Invariant

spatial domain

frequency domain

Charles Cadieu, cadieu@berkeley.edu
Learned Spatial Invariant

spatial domain

frequency domain

Charles Cadieu, cadieu@berkeley.edu
Learned Spatial Invariants

Charles Cadieu, cadieu@berkeley.edu
Feedback

\[ P(\phi_i(t) | \phi_i(t-1), w(t)) \propto e^k \cos(\phi_i(t) - \phi_i(t-1) - [D\ w]_i) \]
Image Denoising: testing an image model
Original | Noisy | Denoised

SNR = -2.9 | SNR = 6.7

Compare to Wiener, SNR = 3.9

Charles Cadieu, cadieu@berkeley.edu
Denoising Movies

Mean SNR of Noisy Images = -2.0

Charles Cadieu, cadieu@berkeley.edu

Wednesday, August 11, 2010
Conclusions

We have,

- Motivated models the produce interpretations of the visual world,
- Learned transformational and spatial invariants from the natural world, and
- Shown how the model improves the interpretation of ambiguous inputs.
Thanks to

Bruno Olshausen

Redwood Center for Theoretical Neuroscience

NSF support
Extra Slides...
Feedback Changes
First Layer Response

With Feedback
Without Feedback

amplitude coefficient

phase coefficient

time frame

time frame
Feedback Changes
First Layer Response

$a_{373}$

$a_{143}$

$\phi_{373}$

$\phi_{143}$