Neural Networks Supporting Persistent Percepts

Shaul Druckmann and Mitya Chklovskii
Janelia Farm Research Campus

NIPS 2010
Working Memory
persistent representation from transient stimuli
Working Memory
persistent representation from transient stimuli
Traditional Explanation:
Constant Percept by Constant Activity
Traditional Explanation:
Constant Percept by Constant Activity

Firing rate

Time

Monday, December 20, 2010
Traditional Explanation:
Constant Percept by Constant Activity

Firing rate

Present Img. 1

Time

Present Img. 2
Traditional Explanation:
Constant Percept by Constant Activity

Firing rate

Time

Present Img. 1 Present Img. 2

Monday, December 20, 2010
Traditional Explanation:
Constant Percept by Constant Activity

Firing rate

Present Img. 1  Time  Present Img. 2

Fuster and Alexander (1971)
Problem: Time-invariant
Activity is rare
Problem: Time-invariant Activity is rare

Brody, alii, et Romo, 2003
Problem: Time-invariant Activity is rare

Brody, alii, et Romo, 2003
Problem: Time-invariant Activity is rare

Can time-variant neuronal activity represent time invariant percepts?

Traditional answer: No!
Linear encoding with orth. basis: persistent percepts → persistent activity
Linear encoding with orth. basis:
persistent percepts $\rightarrow$ persistent activity

$$= a_1 \times \begin{array} \text{image} \end{array} + a_2 \times \begin{array} \text{image} \end{array} + a_3 \times \begin{array} \text{image} \end{array} + a_4 \times \begin{array} \text{image} \end{array} + \ldots$$
Linear encoding with orth. basis:
persistent percepts → persistent activity

\[ s = D a \]
Linear encoding with orth. basis: persistent percepts $\rightarrow$ persistent activity

$$s = D\alpha$$

Persistent percept: $$\frac{ds}{dt} = 0$$
Linear encoding with orth. basis: persistent percepts → persistent activity

\[ s = D \alpha \]

Persistent percept: \[ \frac{ds}{dt} = 0 \]
Linear encoding with orth. basis:
persistent percepts $\rightarrow$ persistent activity

$$s = a_1 \times + a_2 \times + a_3 \times + a_4 \times + ...$$

$$s = D a$$

Persistent percept: $$\frac{ds}{dt} = 0 \quad \quad \rightarrow \quad \quad$$ Persistent activity: $$\frac{da}{dt} = 0$$

Monday, December 20, 2010
Thalamus-Cortex Divergence

Input structure to Cortex is the Thalamus

Thalamus (LGN)
Thalamus-Cortex Divergence

Input structure to Cortex is the Thalamus

Number of cortical neurons much greater than the number of thalamic input neurons

Thalamus (LGN)

Cortex (V1)

Cortex uses a non-orthogonal (over-complete) representation
Linear Encoding in Overcomplete Representation?

\[ s = a_1 \times + a_2 \times + a_3 \times + a_4 \times + \ldots \]

\[
\begin{bmatrix}
4 \\
8 \\
6 \\
3
\end{bmatrix}
\quad = 
\begin{bmatrix}
1 \\
3
\end{bmatrix}
\times
\begin{bmatrix}
\text{D} \\
\text{a}
\end{bmatrix}
\]

\[ s = Da \]

Persistent percept: \[ \frac{ds}{dt} = 0 \]
Linear Encoding in Overcomplete Representation?

\[ s = a_1 \times + a_2 \times + a_3 \times + a_4 \times + \ldots \]

\[ s = D a \]

Persistent percept: \( \frac{ds}{dt} = 0 \implies ? \]
Over-complete Representation is Non-unique
Over-complete Representation is Non-unique

Stimulus dimension: 2

Number of neurons (activity dimension): 3
Over-complete Representation is Non-unique

Stimulus dimension: 2

Number of neurons (activity dimension): 3
Over-complete Representation is Non-unique

Stimulus dimension: 2

Number of neurons (activity dimension): 3

Monday, December 20, 2010
Over-complete Representation is Non-unique
Over-complete Representation is Non-unique
Over-complete Representation is Non-unique
Freedom in Representation in an Over-complete Frame
Freedom in Representation in an Over-complete Frame
Freedom in Representation in an Over-complete Frame
Freedom in Representation in an Over-complete Frame

Firing rate $a_1$

Time

$-a_2$

$-a_3$


Monday, December 20, 2010
Time-variant neuronal activity can represent a time invariant percept.
Lateral connectivity maintains persistency

s: stimulus, a: activity
D: dictionary (feature vectors), L: lateral connections
Lateral connectivity maintains persistency

\[ s = Da \]

Linear encoding

- **s**: stimulus, **a**: activity
- **D**: dictionary (feature vectors), **L**: lateral connections

Monday, December 20, 2010
Lateral connectivity maintains persistency

\[ s = Da \]

\[ \dot{a} = -a + La \]

Linear encoding
Rate dynamics

\( s \): stimulus, \( a \): activity
\( D \): dictionary (feature vectors), \( L \): lateral connections
Lateral connectivity maintains persistency

\[ s = Da \]
\[ \dot{a} = -a + La \]
\[ \dot{s} = D \dot{a} = 0 = D(-a + La) \]

**Linear encoding**

**Rate dynamics**

**Persistency**

\( s: \text{stimulus}, \ a: \text{activity} \)

\( D: \text{dictionary (feature vectors)}, \ L: \text{lateral connections} \)
Lateral connectivity maintains persistency

\[ s = Da \]  
\[ \dot{a} = -a + La \]  
\[ \dot{s} = D\dot{a} = 0 = D(-a + La) \]  
\[ Da = DLa \]

Linear encoding  
Rate dynamics  
Persistency  
If to hold for all \( a \)

**s**: stimulus, **a**: activity  
**D**: dictionary (feature vectors), **L**: lateral connections
Lateral connectivity maintains persistency

\[ s = Da \]

\[ \dot{a} = -a + La \]

\[ \dot{s} = D\dot{a} = 0 = D(-a + La) \]

\[ Da = DLa \]

\[ D = DL \]

Linear encoding
Rate dynamics
Persistency
If to hold for all \( a \)
Family of Solutions

\textbf{s}: stimulus, \( a \): activity

\textbf{D}: dictionary (feature vectors), \( L \): lateral connections
Lateral connectivity maintains persistency

\[ s = Da \]  \hspace{2cm} \text{Linear encoding}

\[ \dot{a} = -a + La \]  \hspace{2cm} \text{Rate dynamics}

\[ \dot{s} = D\dot{a} = 0 = D(-a + La) \]  \hspace{2cm} \text{Persistency}

\[ Da = DLa \]  \hspace{2cm} \text{If to hold for all } a

\[ D = DL \]  \hspace{2cm} \text{Family of Solutions}

\[ L = I \]  \hspace{2cm} \text{Trivial solution}

s: stimulus, a: activity
D: dictionary (feature vectors), L: lateral connections
Lateral connectivity maintains persistency

\[ s = Da \quad \text{Linear encoding} \]
\[ \dot{a} = -a + La \quad \text{Rate dynamics} \]

\[ D = DL \quad \text{Family of Solutions} \]

s: stimulus, a: activity
D: dictionary (feature vectors), L: lateral connections
Our Solution: sparse solution

\[ D = DL \]
Our Solution: sparse solution

\[ D = DL \]
Our Solution: sparse solution

Entries in $L$ represent synaptic connections

$$D = DL$$
Our Solution: sparse solution

Entries in L represent synaptic connections

We pick the most economic solution, in terms of the resources taken up by synapses

\[ D = D \cdot L \]
Our Solution: sparse solution

Entries in L represent synaptic connections

We pick the most economic solution, in terms of the resources taken up by synapses

$$\min_L: \left[ (D - DL)^2 + \lambda |L|_1 \right]$$

Reconstruction Error \hspace{1cm} Sparsity

$$D = DL$$

$$D = \begin{bmatrix} D \\ L \end{bmatrix}$$

Monday, December 20, 2010
Neurons compensate for their changing representation by modifying post-synaptic neuronal activity.
Neurons compensate for their changing representation by modifying post-synaptic neuronal activity

\[ D = DL \]
Neurons compensate for their changing representation by modifying post-synaptic neuronal activity

\[ D = DL \]
Neurons compensate for their changing representation by modifying post-synaptic neuronal activity

\[ D = DL \]

Sum of outgoing synapses times post-syn. receptive field equals neuron’s own receptive field
Neurons compensate for their changing representation by modifying post-synaptic neuronal activity.

\[ D = DL \]

Sum of outgoing synapses times post-syn. receptive field equals neuron’s own receptive field.
Neurons compensate for their changing representation by modifying post-synaptic neuronal activity.

\[ D = DL \]

Sum of outgoing synapses times post-syn. receptive field equals neuron’s own receptive field.
REceptive Field RE-combination (REFIRE) guarantees persistent percepts

Sum of outgoing synapses times post-syn. receptive field equals neuron’s own receptive field

\[ D = DL \]
REFIRE with V1 receptive fields
REFIRE with V1 receptive fields

Receptive fields after: Olshausen and Field, 1997
REFIRE with V1 receptive fields

\[ D = DL \]

Receptive fields after: Olshausen and Field, 1997
REFIRE with V1 receptive fields

\[ D = DL \]

Receptive fields after: Olshausen and Field, 1997
REFIRE with V1 receptive fields

\[ D = DL \]

Receptive fields after: Olshausen and Field, 1997
REFIRE with V1 receptive fields

\[ D = DL \]

Receptive fields after: Olshausen and Field, 1997

Monday, December 20, 2010
Network Structure
Network Structure
Numerical validation: Percepts are Persistent
Neuron 1

Neuron 2

Neuron 3

Neuron 4

Neuron 5

Neuron 6

Network
Synaptic weight distribution matches Experiments
Synaptic weight distribution matches Experiments

Song, Sjostrom, Reigl, Nelson, Chklovskii (2005)
Synaptic weight distribution matches Experiments

Song, Sjostrom, Reigl, Nelson, Chklovskii (2005)
Network Motifs match Experiments
Network Motifs match Experiments

2 Neuron motifs:
Network Motifs match Experiments

2 Neuron motifs:

A → B

A ← B

A → B

A → B

P_{precip}
Network Motifs match Experiments

2 Neuron motifs:

A → B
B → A

In cortex and in REFIRE network: \( P_{\text{recip}} > p \times p \)
Network Motifs match Experiments

2 Neuron motifs:

In cortex and in REFIRE network: $P_{\text{recip}} > p \times p$

3 Neuron motifs:
Network Motifs match Experiments

2 Neuron motifs:

\[ \begin{align*}
&\text{A} \quad p \quad \text{B} \\
&\text{A} \quad p \quad \text{B} \\
&\text{A} \quad P_{\text{recip}} \quad \text{B}
\end{align*}\]

In cortex and in REFIRE network: \( P_{\text{recip}} > p \times p \)

3 Neuron motifs:
Network Motifs match Experiments

2 Neuron motifs:

\[ \text{2 Neuron motifs:} \]

\[ \begin{align*}
\text{A} & \rightarrow \text{B} \\
\text{p} & \\
\text{A} & \leftarrow \text{B} \\
\text{p} & \\
\end{align*} \]

In cortex and in REFIRE network: \( P_{\text{recip}} > p \times p \)

3 Neuron motifs:

\[ \text{3 Neuron motifs:} \]

\[ \begin{align*}
\text{Over expression of reciprocal motifs} \\
\end{align*} \]

Monday, December 20, 2010
Summary
Can time-variant neuronal activity represent time invariant percepts?
Summary

Can time-variant neuronal activity represent time invariant percepts?

Yes. But not in an orthogonal basis!
Can time-variant neuronal activity represent time invariant percepts?

Yes. But not in an orthogonal basis!

We propose a specific form of a network that supports time invariant percepts with time-variant neuronal activity: REFIRE network
Can time-variant neuronal activity represent time invariant percepts?

Yes. But not in an orthogonal basis!

We propose a specific form of a network that supports time invariant percepts with time-variant neuronal activity: REFIRE network

This network qualitatively matches known statistical properties of cortical networks
Thanks
Thanks

Mitya Chklovskii
Thanks

Mitya Chklovskii

Shiv Vitaladevuni, Tao Hu, William Katz, Juan Nunez-Iglesias, Arjun Bharioke, Anatoli Grinspan and Lav Varshney

Frank Midgley

and thank you for your attention!
Poster: T9

druckmanns@janelia.hhmi.org
Similar Orientations, More Connections
Similar Orientations, More Connections
Similar Orientations, More Connections
Similar Orientations, More Connections
Similar Orientations, More Connections
Similar Orientations, More Connections
Similar Orientations, More Connections

Monday, December 20, 2010
Similar Orientations, More Connections
Similar Orientations, More Connections
Similar Orientations, More Connections

Youssef et al. (1999)
Similar Orientations, More Connections

Youssef et al. (1999)
Similar Orientations, More Connections

Youssef et al. (1999)
Similar Orientations, More Connections

Youssef et al. (1999)

Monday, December 20, 2010
Similar Orientations, More Connections

Youssef et al. (1999)

Monday, December 20, 2010
Similar Orientations, More Connections

Youssef et al. (1999)

Monday, December 20, 2010
Persistence across Cortex
Persistence across Cortex

Hernandez, alii, et Romo (2010)