ISSA 2017 MRI hands-on project (Call for your ideas)

Aim: introducing MRI experiments and modeling analyses

Experiment 1: building a semantic model of the brain

Experiment 2: decoding brain activity evoked under diverse subjective experiences

Call you your solid/radical/enjoyable inputs!

e.g., attending to humans vs. attending to vehicles (Çukur et al., 2013)

- Voluntary actions vs. Guided actions
- Focusing vs. Wandering
- Epoché vs. Objective observation
- …

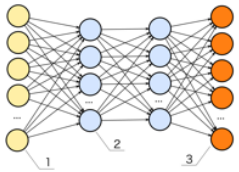
Talk summary

1. Visual spatiotemporal representation

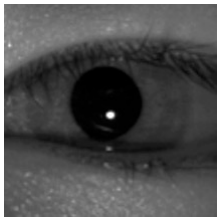
2. Visual category representation

3. Language model representation

4. Using “AI” to decode brain



5. Eye movement-invariant representation



Modeling semantic representation in the brain

Visual experience

Natural language description

A. The man, bearded and a brown scarf wrapped around his neck, is looking this way. His mouth is open, his teeth are straight, and his fringe is long.

B. This is some sort of scene from a movie. I was thinking it could have been a western, but could see something like a large bridge faintly in the background, so the period isn't clear.

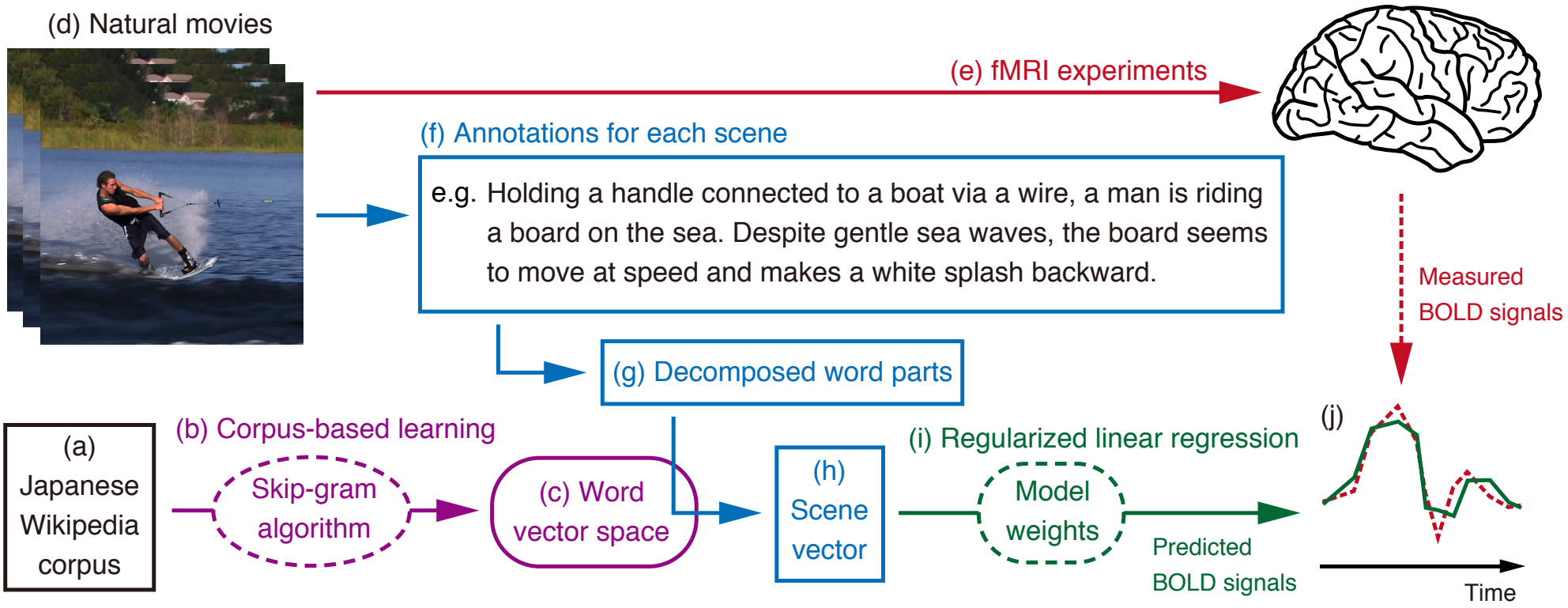
C. The man in the front seems to be a defender of justice, there to save the poor. The people in the back must be supporting him.

D. I wonder if it's really a scene from a movie. The men seem to be blue-collar workers. They seem a bit sweaty and dirty, but in high spirits.

E. The man with the thick, slightly wild-looking beard is on the right, talking while facing this way. The people in the back are watching over him from afar.

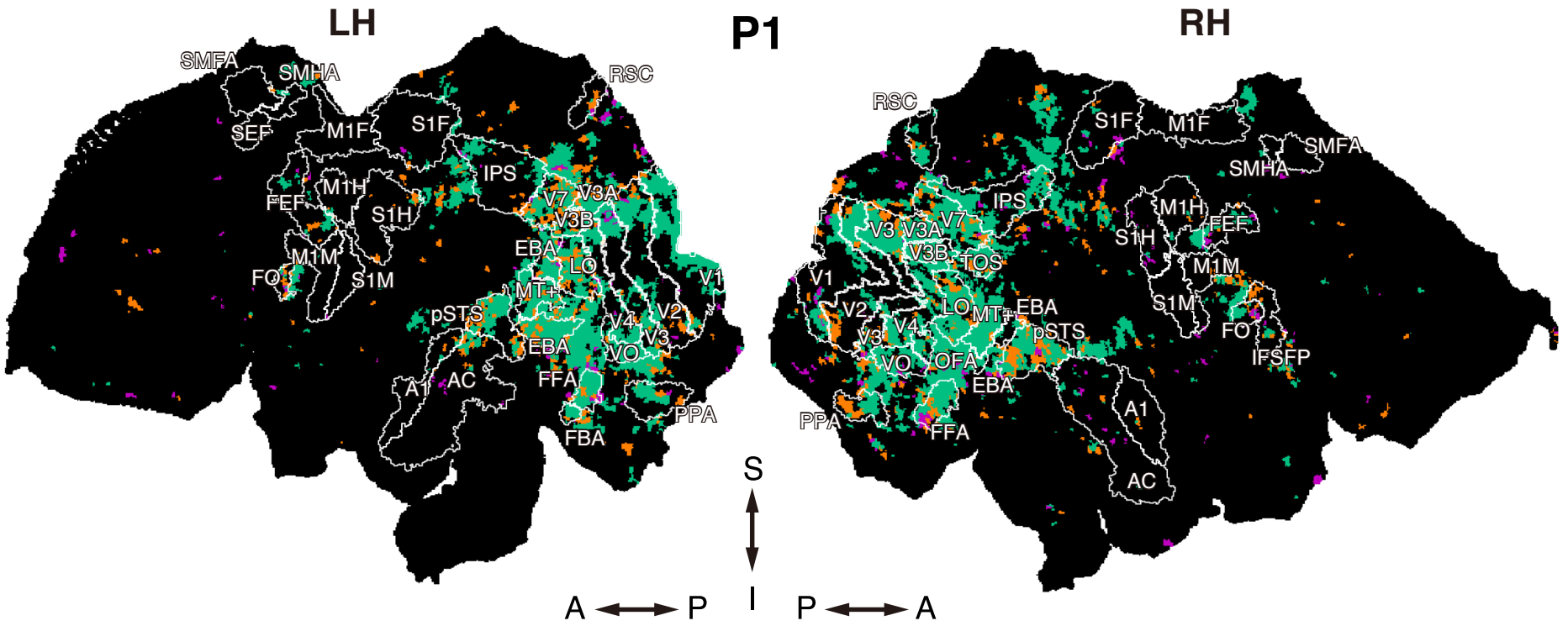
- Context, subjective impression, memory,…
- All perceptual and cognitive contents reflect brain activity; thus should be modelable

Modeling brain activity via language representation



Nishimoto and Kashioka 2015 *patent pending*; Nishida et al., *under review*

The skip-gram model outperforms traditional language models



- LSI: Latent Semantic Indexing (Deerwester et al., 1990 *JASIS*)

  (Brain models: Carota et al., 2017 *Cerebral Cortex*)

- LDA: Latent Dirichlet Allocation (Blei et al., 2003 *JMLR*)

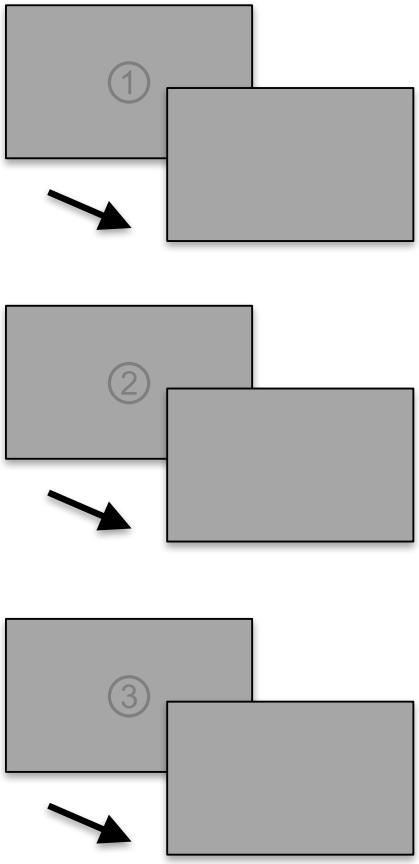
  (Brain models: Stansbury et al., 2013 *Neuron*)

- WordNet (Princeton Univ. 2010)

  (Brain models: Huth et al., 2012 *Neuron*; Horikawa et al., 2013 *Science*)

Toward high-bandwidth vision BMI: Decoding imagined contents

Stimuli



Imagined movie scenes

Decoded contents from brain activity

woman	**yearn**	**young**
senior	wear	intimate
child	meet	childish
girl	get to know	fun

animal	**line up**	**narrow**
feed	run	shallow
rabbit	extend	large
tail	chase	thin

building	**line up**	**new**
removal	replace	old
company	establish	
display	mount	

Miyake et al. *JNNS2016*

cf. Mar. 2017 **Neuralink** (Elon Musk)
invasive BMI for normal people in 10 yr.
Apr. 2017 **FaceBook "Building 8"** (Regina Dugan)
non-invasive BMI for 100 words/min in 2 yr.

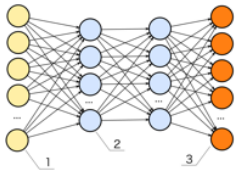
Talk summary

1. Visual spatiotemporal representation

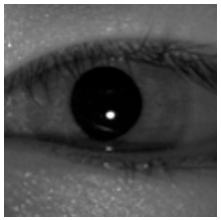
2. Visual category representation

3. Language model representation

4. Using “AI” to decode brain

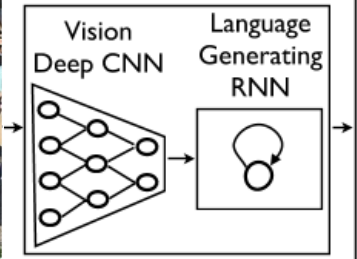


5. Eye movement-invariant representation



Combining “artificial intelligence” and neuroscience

Automatic image caption generator



A group of people shopping at an outdoor market.

There are many vegetables at the fruit stand.

A group of young people playing a game of frisbee.

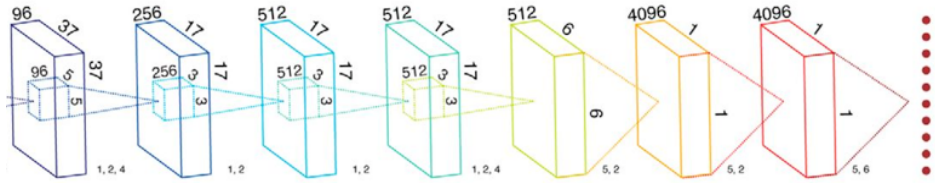
…

A little girl in a pink hat is blowing bubbles.

Vinyals et al., 2015 *CVPR*;
Xu et al., 2015 *ICML*

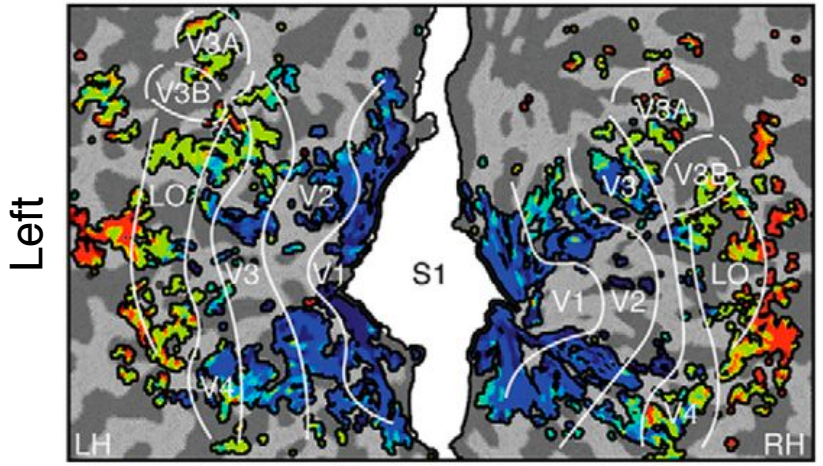
Image → CNN → RNN → Caption

Hierarchical representation of CNN



early → late

Hierarchical representation of the visual cortex



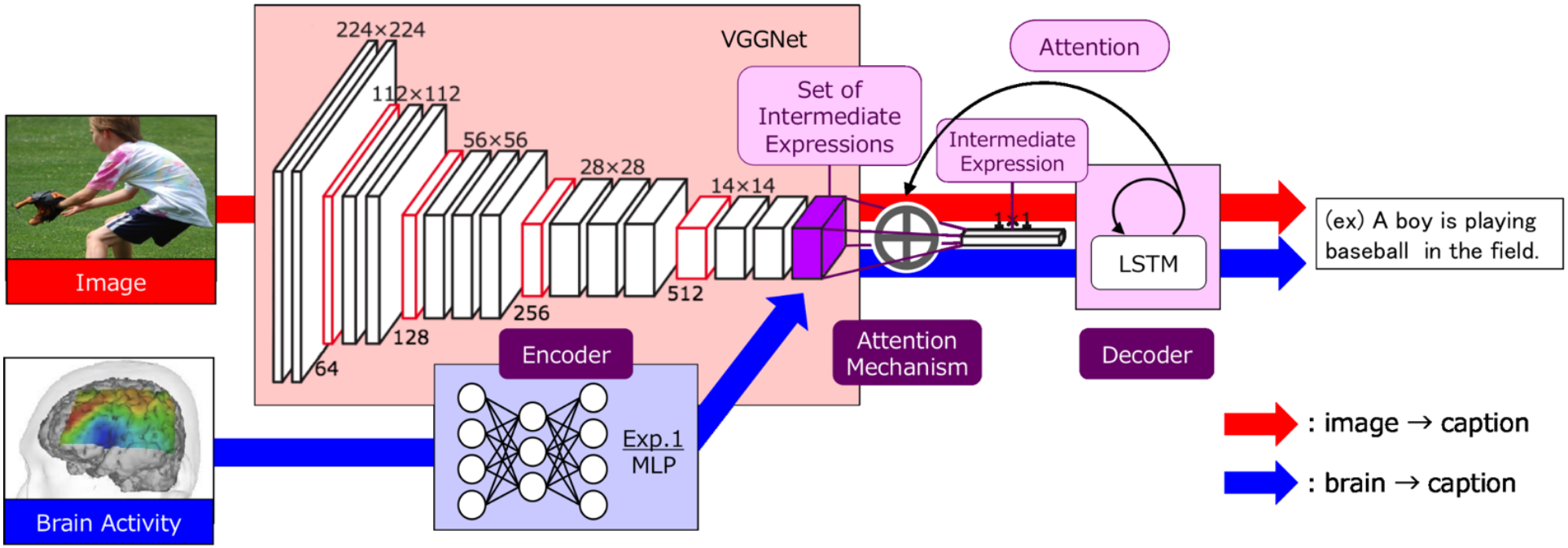
late ← early early → late

Güçlü and van Gerven 2015, 2016;
Yamins et al., 2014;

CNN ≅ Visual Brain

Can we combine the two?

Decoding texts from brain activity



Matsuo et al., *ACL SRW 2016*

Experience	Decoded text from brain activity
	**A group of people standing on the beach.**
	**A man is in the back of an umbrella.**

Talk summary

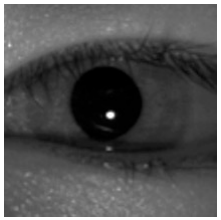
1. Visual spatiotemporal representation

2. Visual category representation

3. Language model representation

4. Using “AI” to decode brain

5. Eye movement-invariant representation

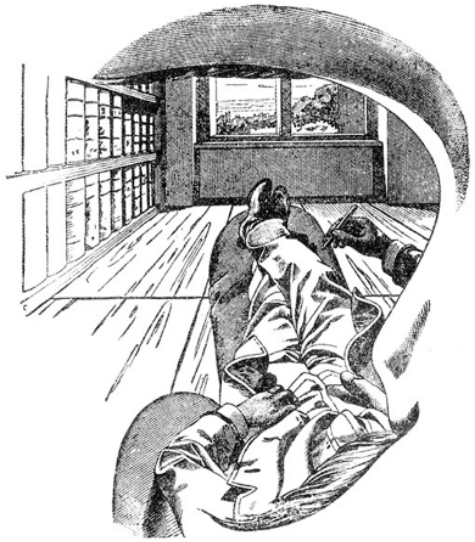


We see a stable visual world while making frequent eye movements

Visual scene

gaze point

Retinal (foveal) input



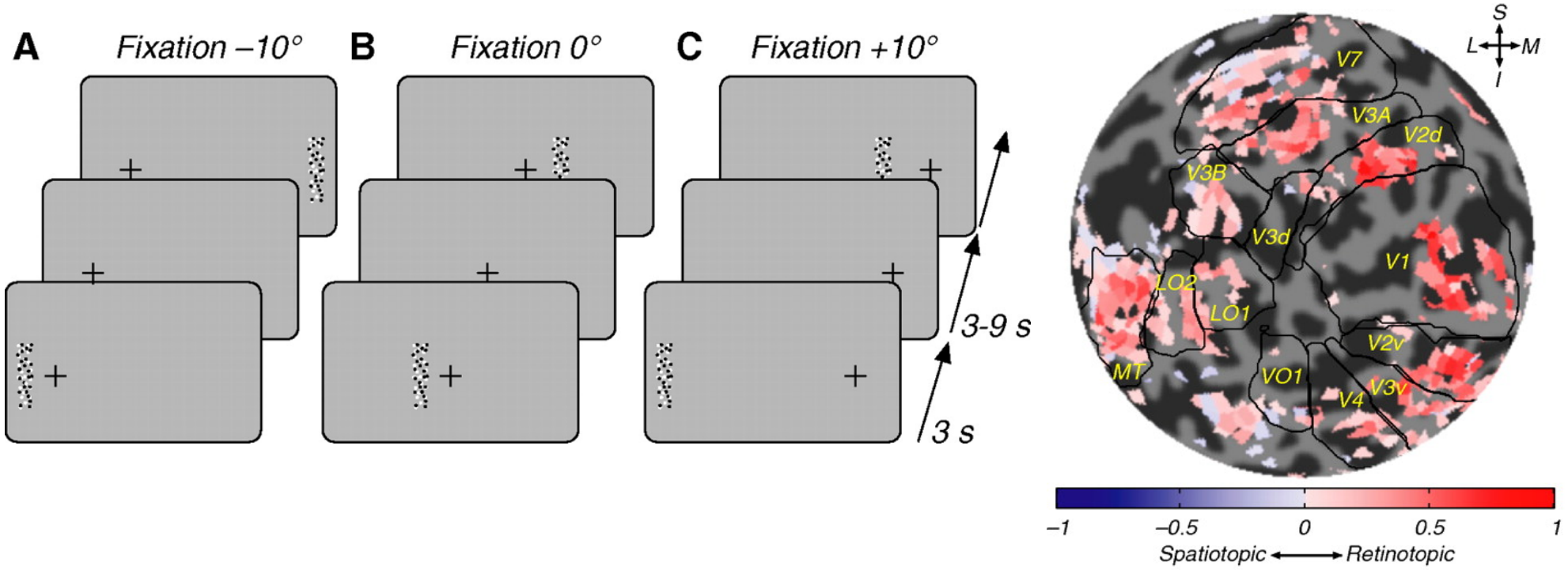
"Illusory" reality

We make saccadic eye movements around 3 times/sec.

Blinks block visual inputs for around 10% of our waking hours.

Question: Are there any visual areas that represent the visual world in an eye movement invariant manner?

Human occipital areas are retinotopic, not spatiotopic

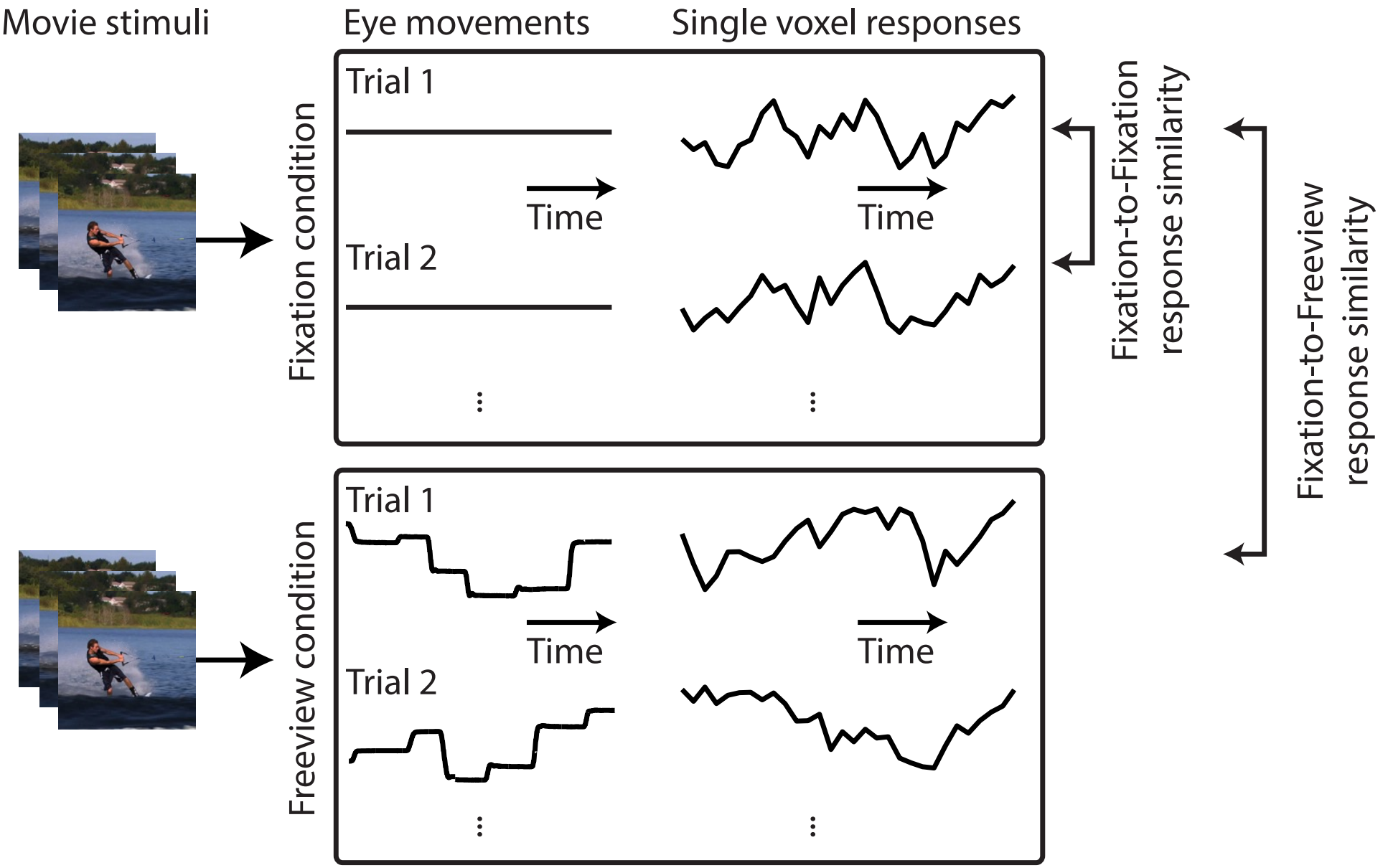


Gardner et al., 2008 *J. Neurosci.*

However…

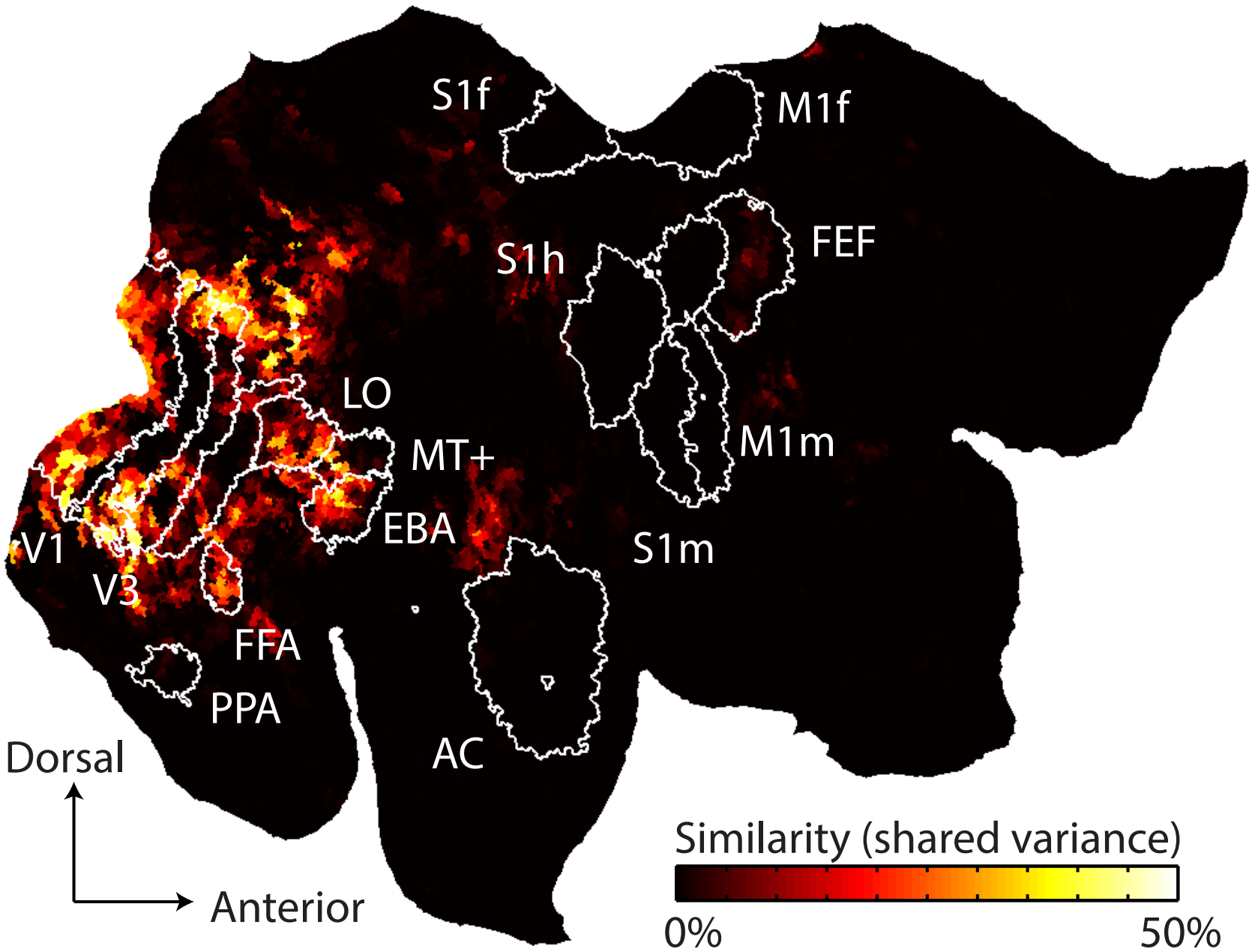
- Fixational position (in)variance ≠ natural eye movement (in)variance
- Simple stimuli do not probe higher areas efficiently

Quantification of eye-movement invariance

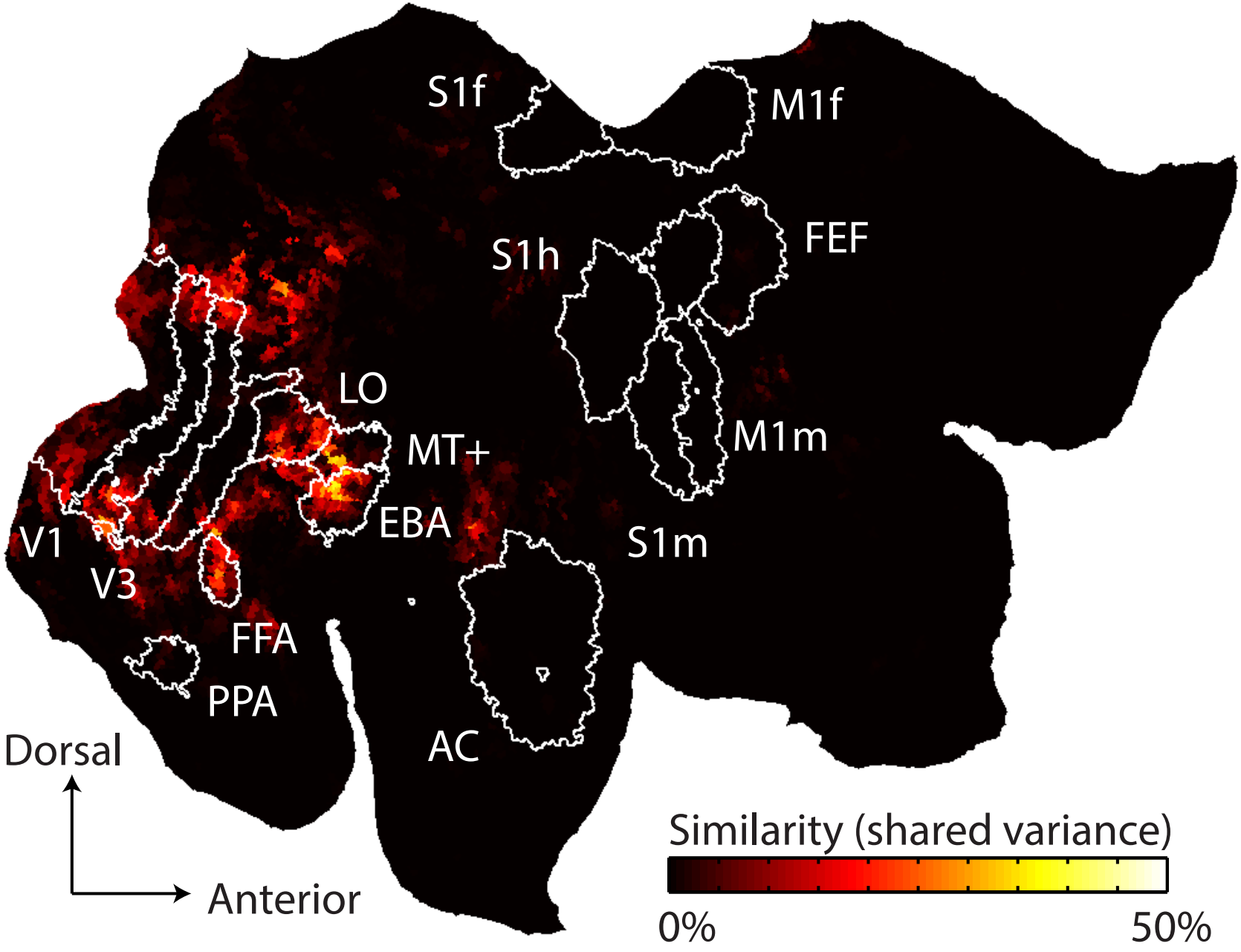


Nishimoto et al., 2017 *Journal of Vision*

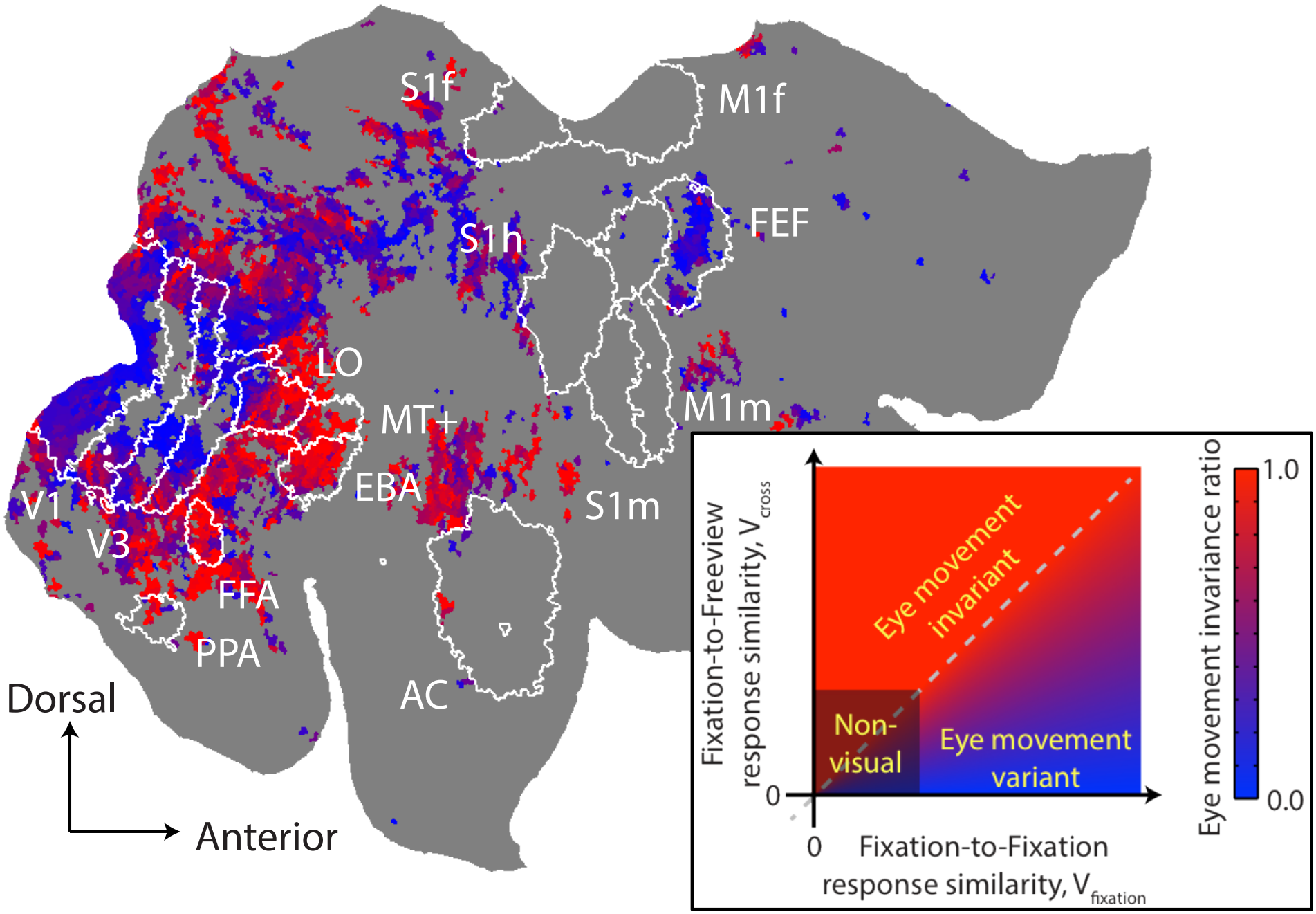
Fixation-to-fixation response similarity ($V_{fixation}$)



Fixation-to-freeview response similarity (V_{cross})



Eye movement invariance ($V_{cross}/V_{fixation}$)



Talk summary

high

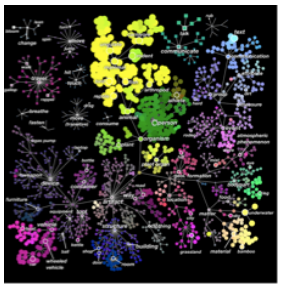
low

Encoding model: p(RIS)

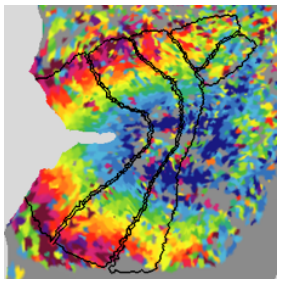
Internal representation
(hypothesis)
Objective features
Subjective features
…

Decoding model: p(SIR)

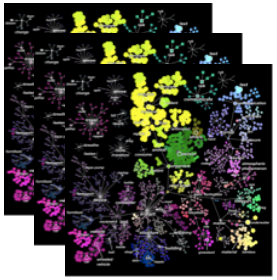
Representation



Mapping



Comparison



Decoding