

b



d



f

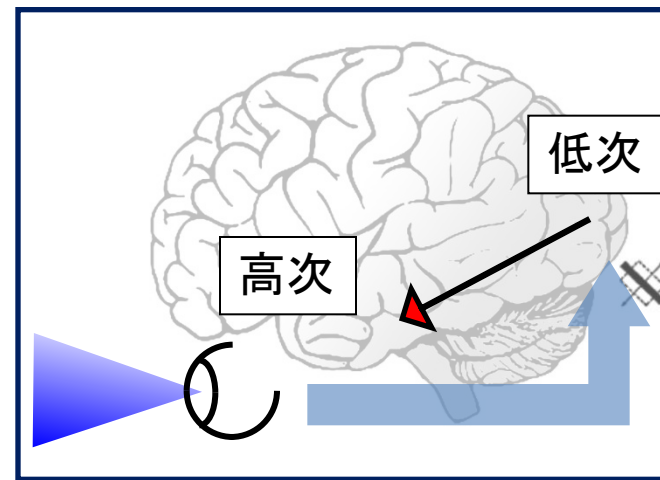
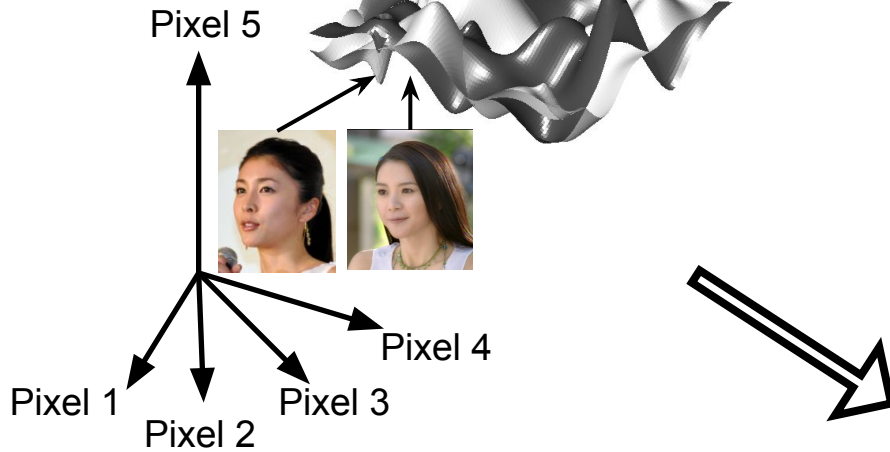
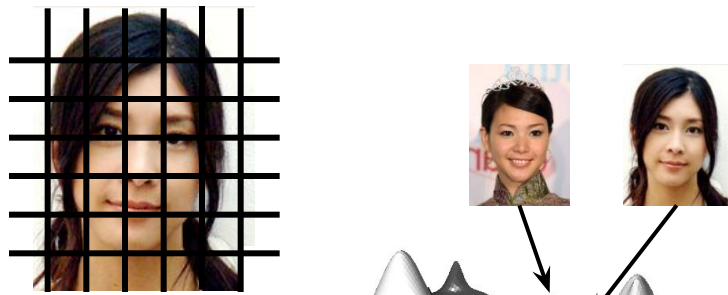


Person
A and B

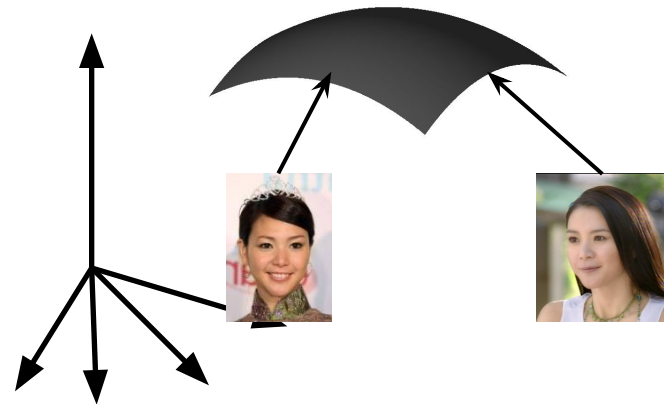


Different views
of person A

網膜上の物体

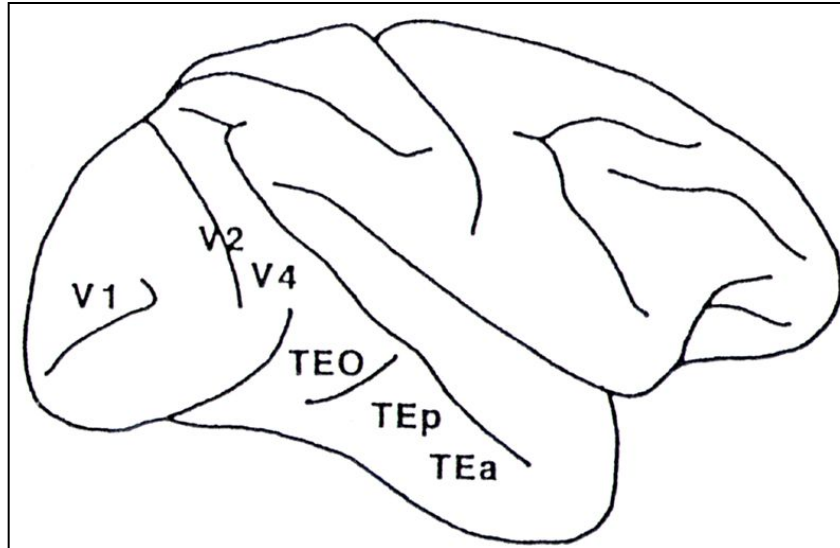


高次視覚野上の物体

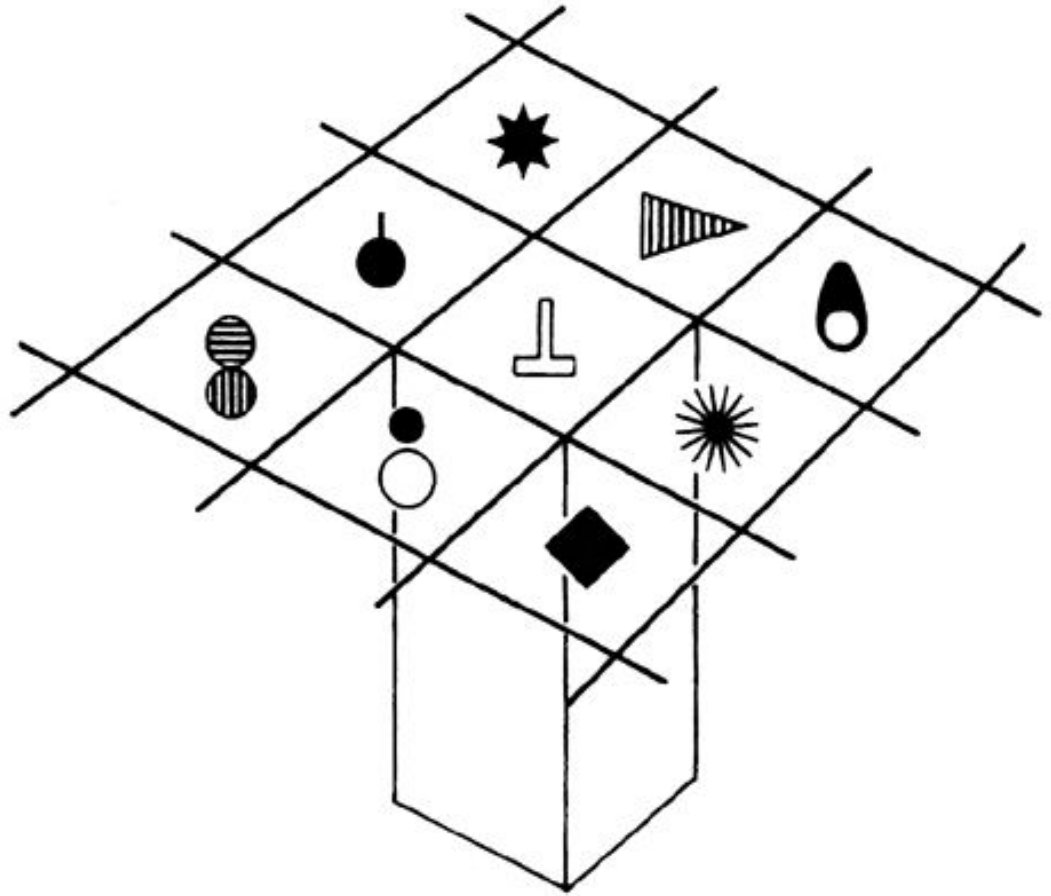
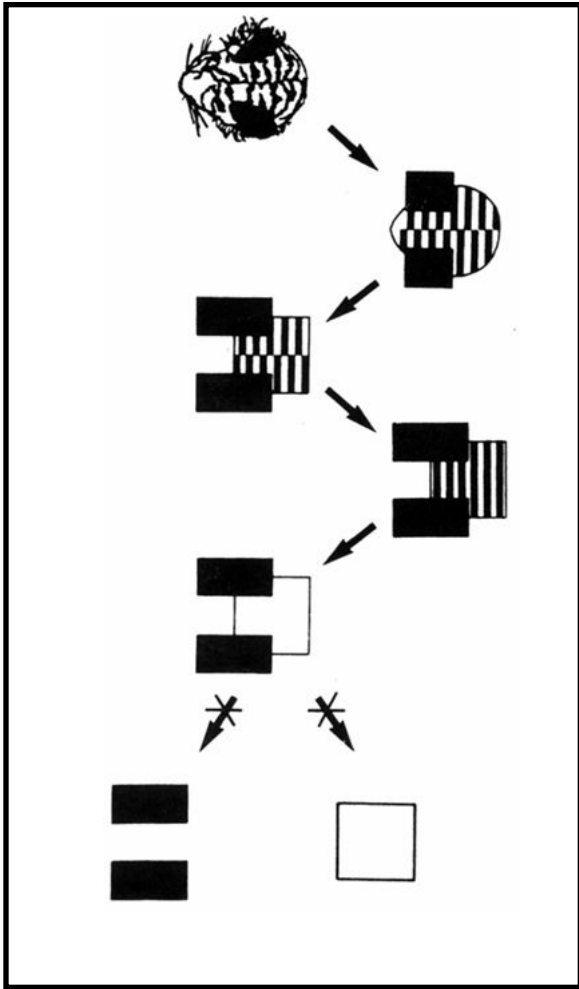


- ✓ 機能単位は何か？一機能を支える構造を知りたい。
- ✓ 機能単位が表現している視覚特徴は？

サル高次視覚野の機能と構造



1. 高次視覚野の機能構造
2. 機能単位が表現する視覚特徴
3. 物体表現のダイナミクス



(Keiji Tanaka, et al. 1993)

細胞1

5.2

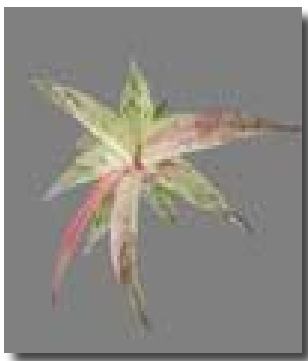
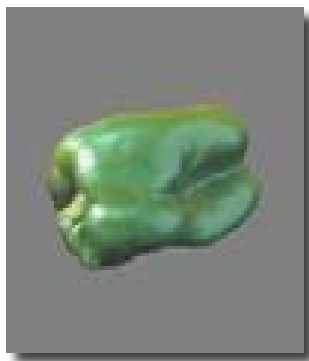
4.0

2.0

1.8

[spike/sec.]

1.7



細胞2

12.5

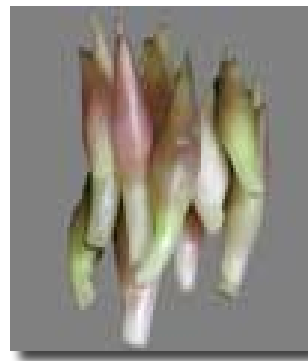
4.0

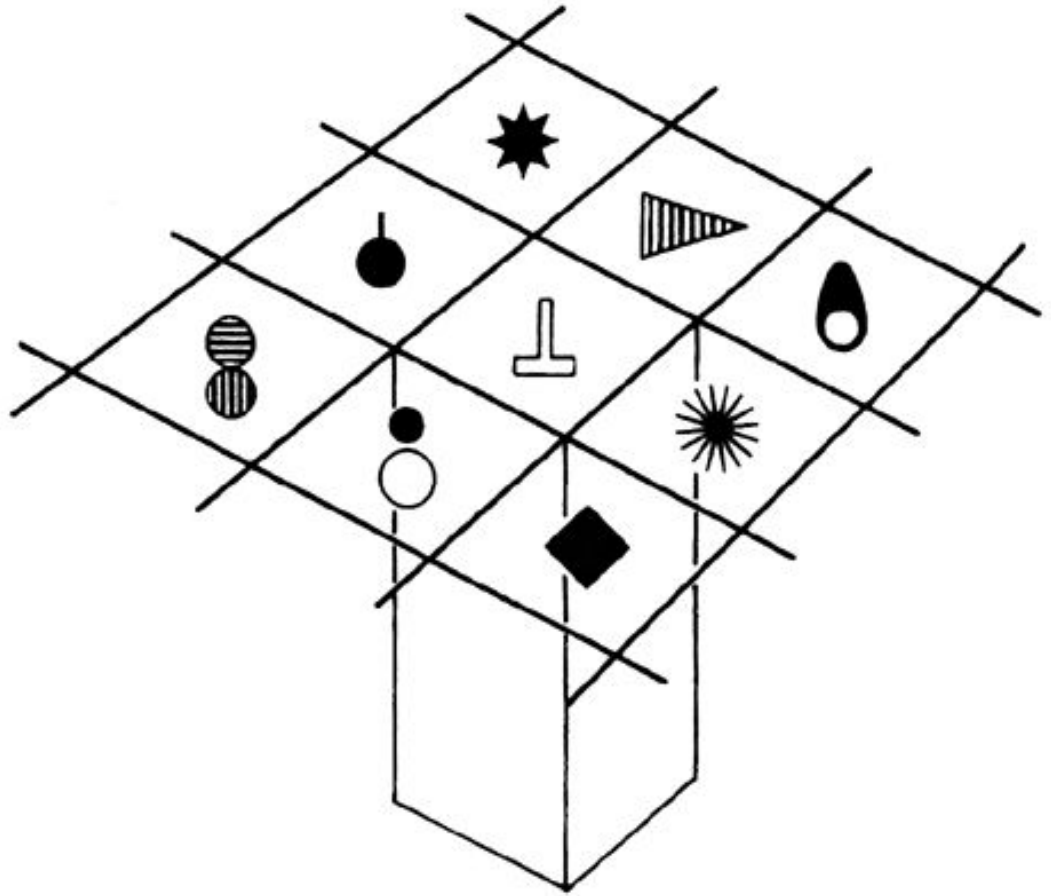
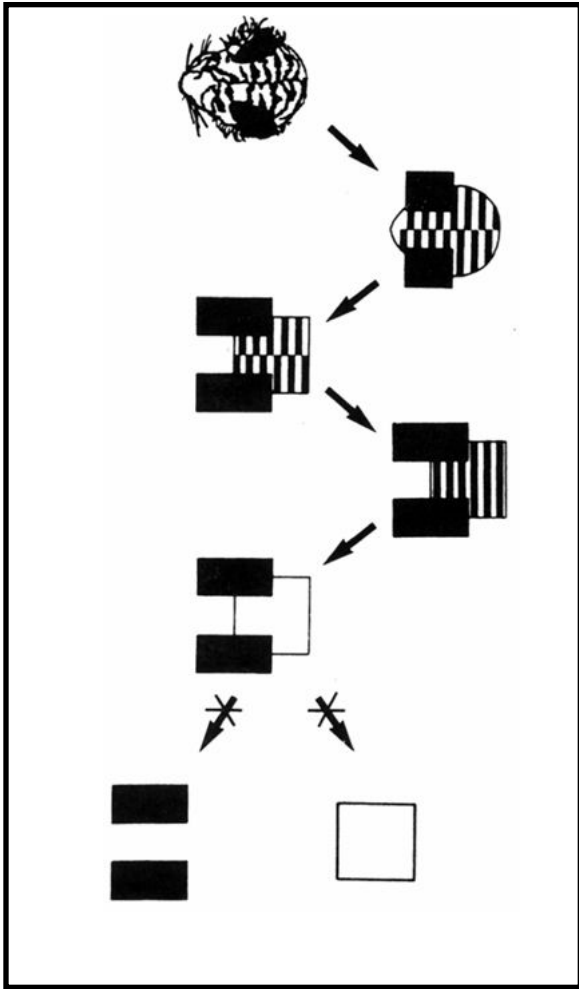
2.8

2.7

[spike/sec.]

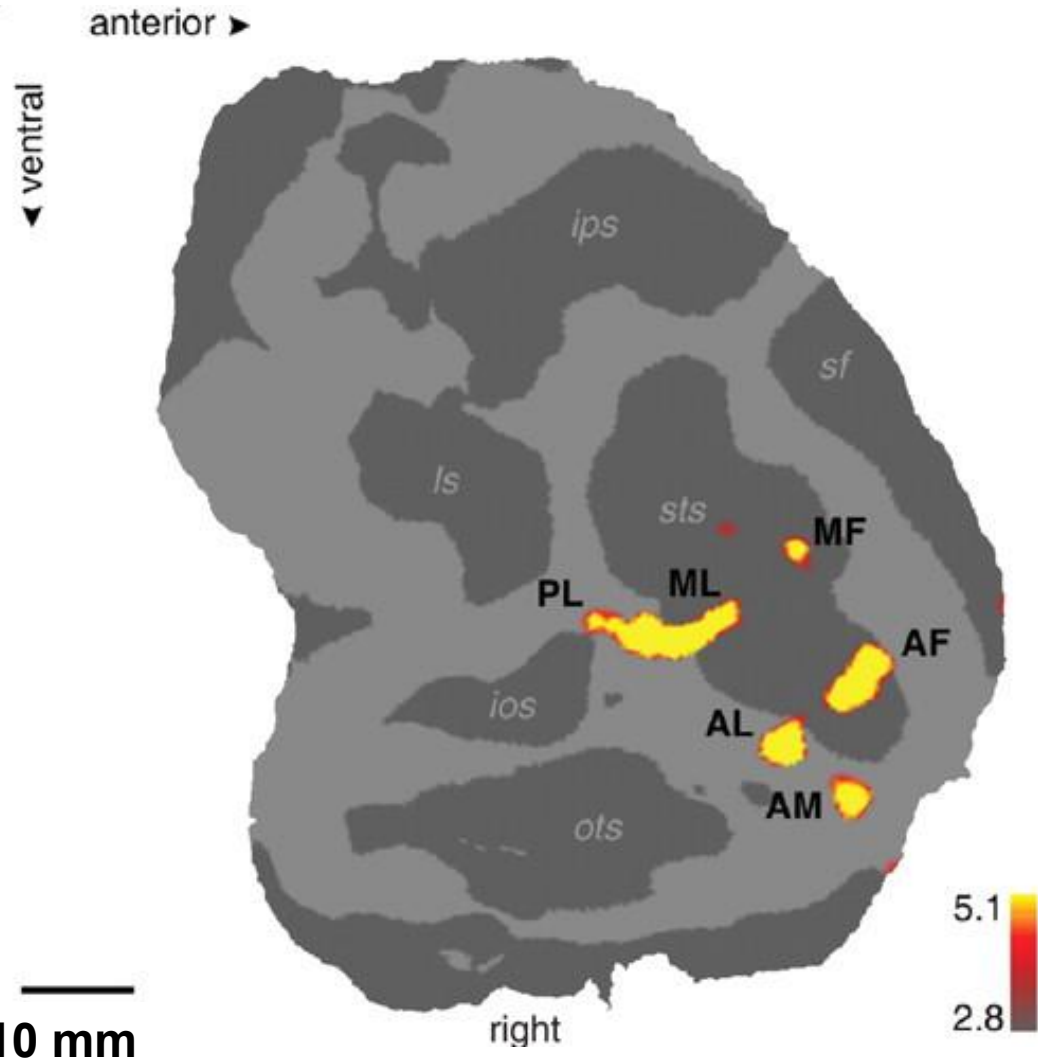
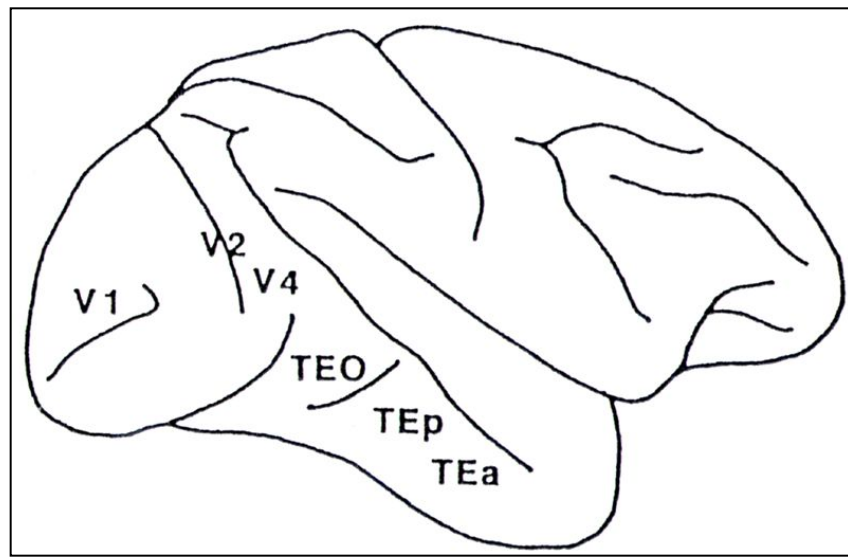
1.7

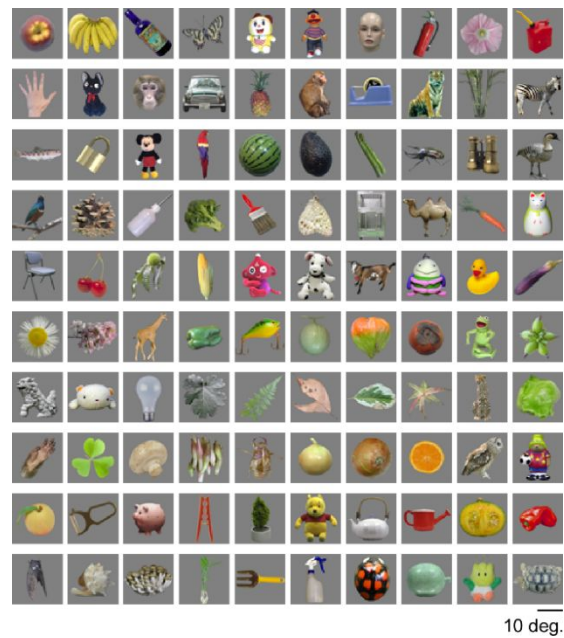
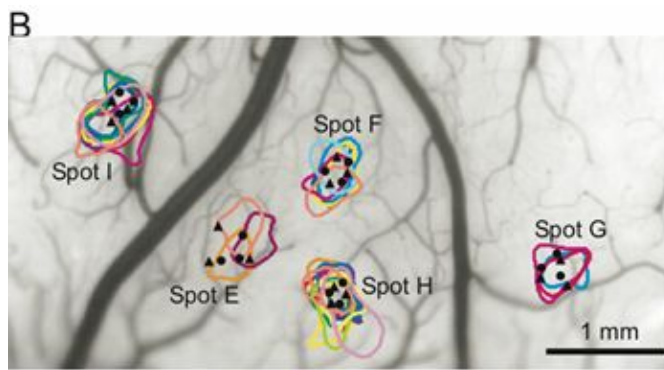
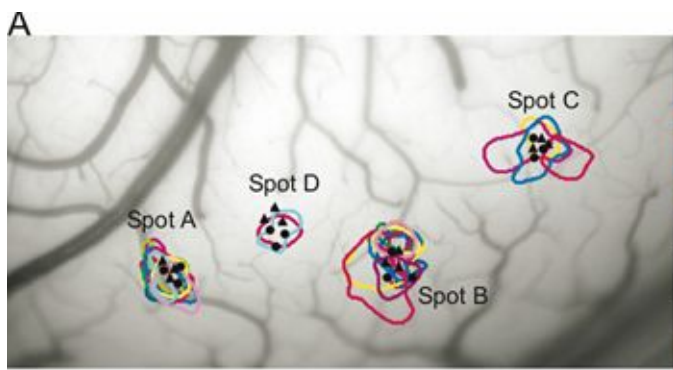




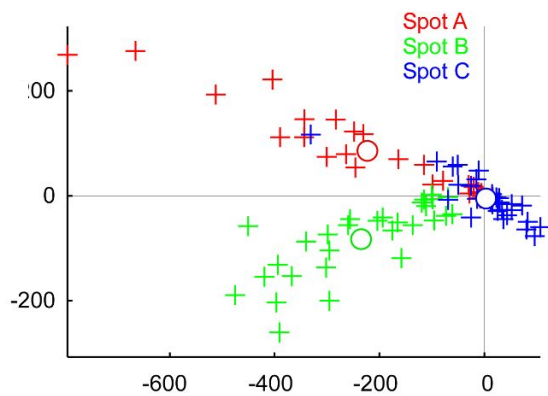
(Keiji Tanaka, et al. 1993)

顔と顔以外の物体

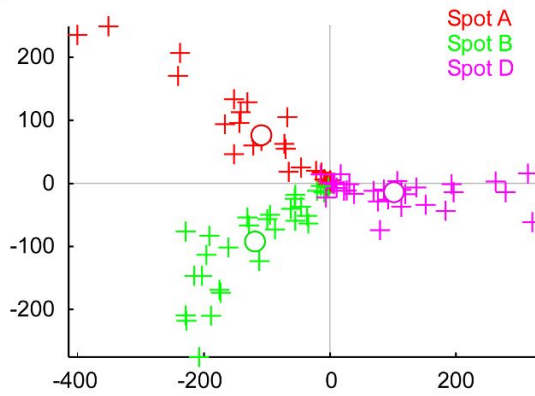




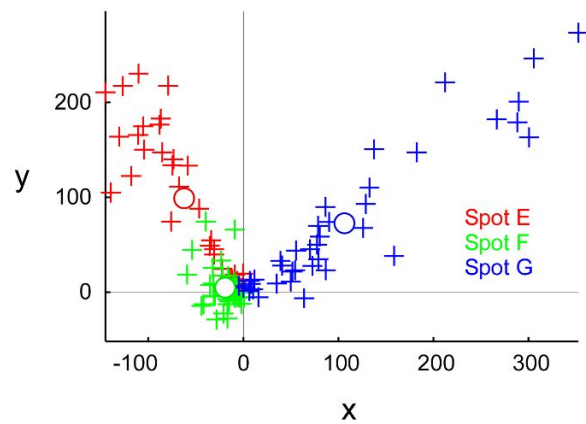
A



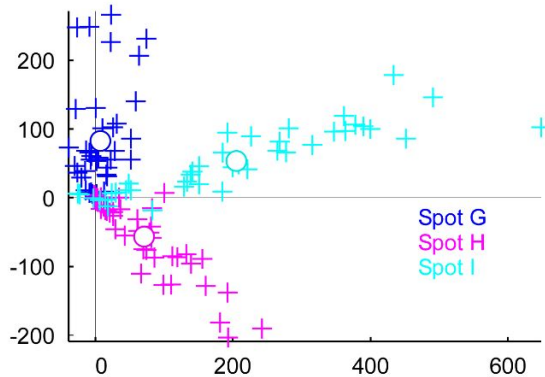
B



C

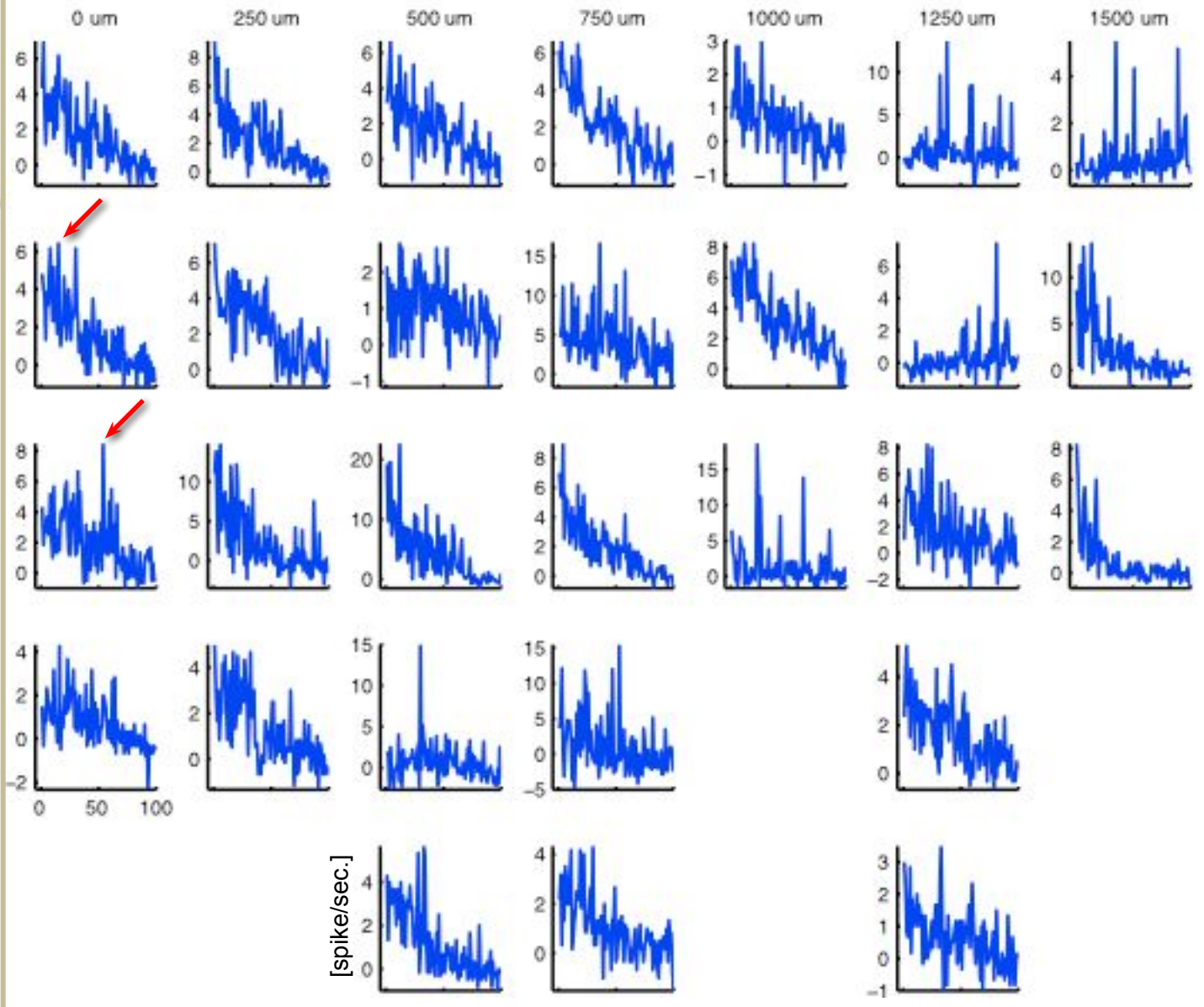
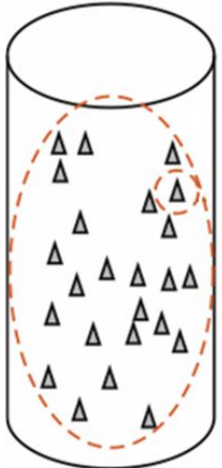
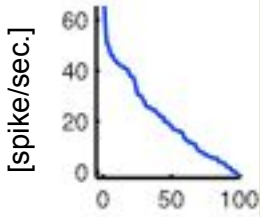


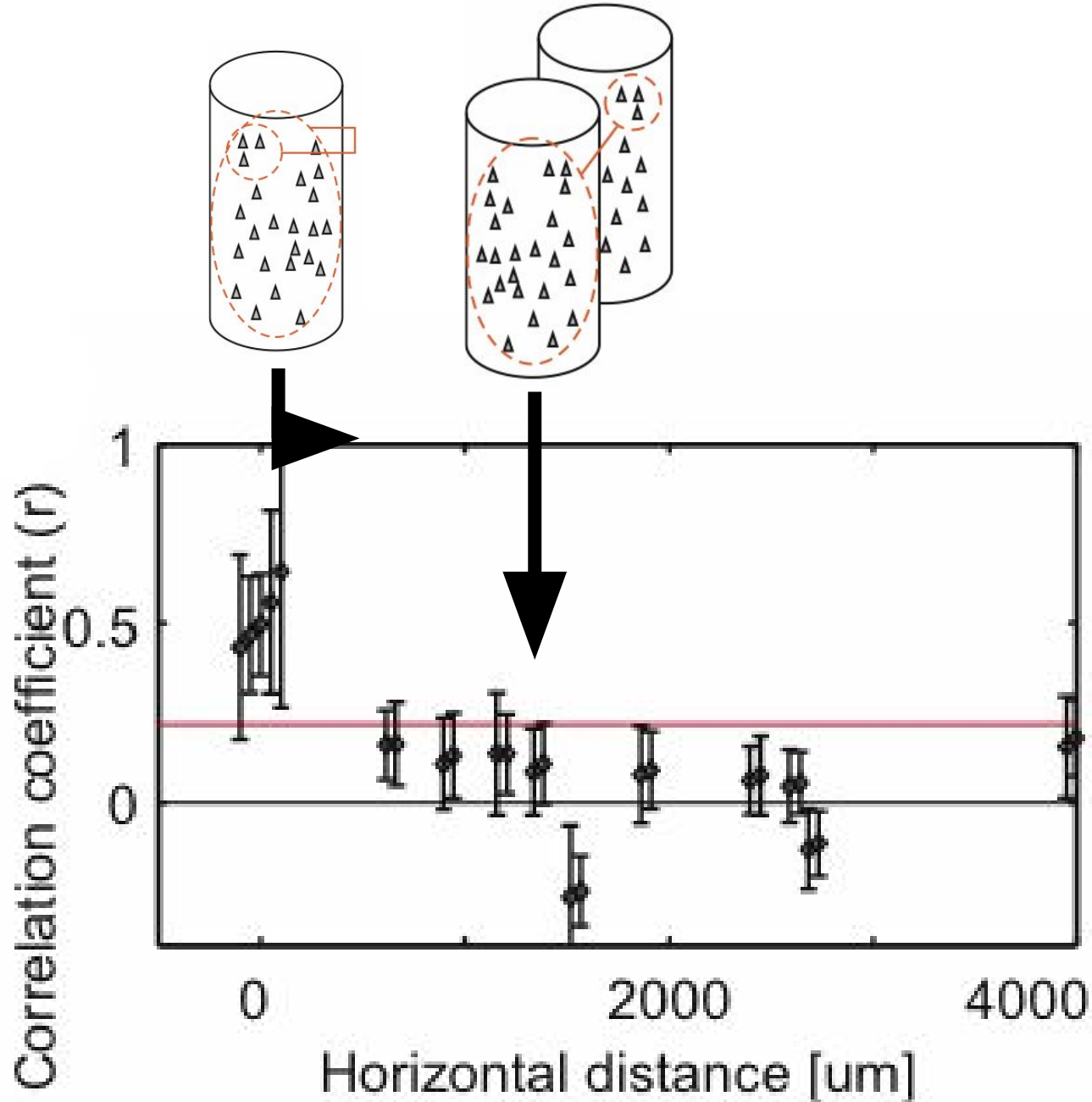
D



Averaged MUAs

Single cells

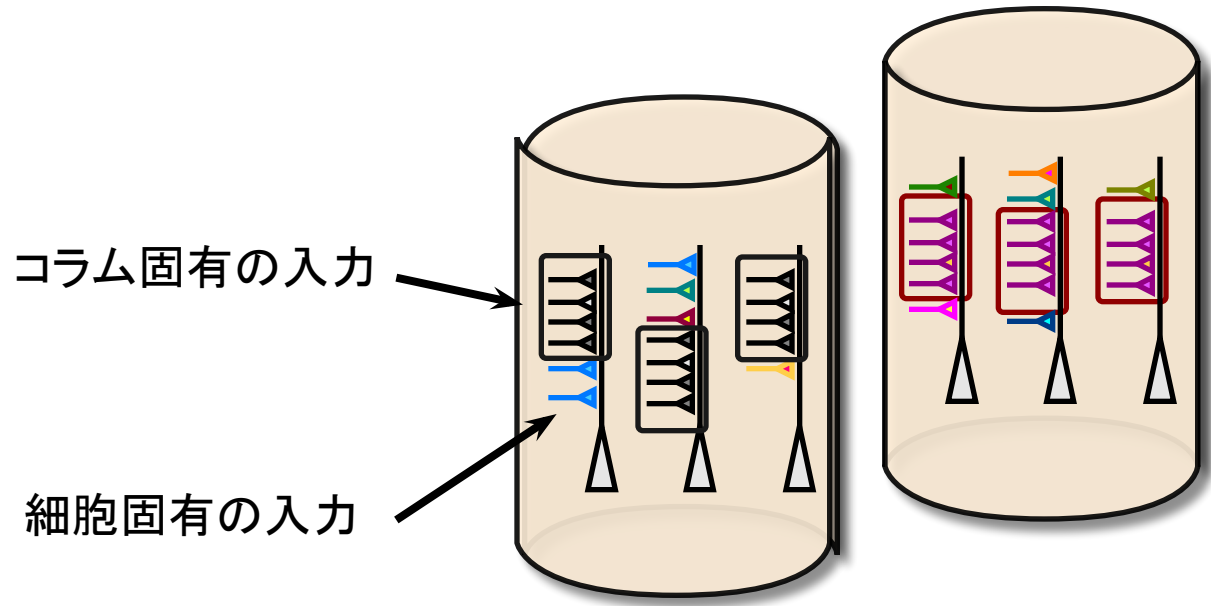


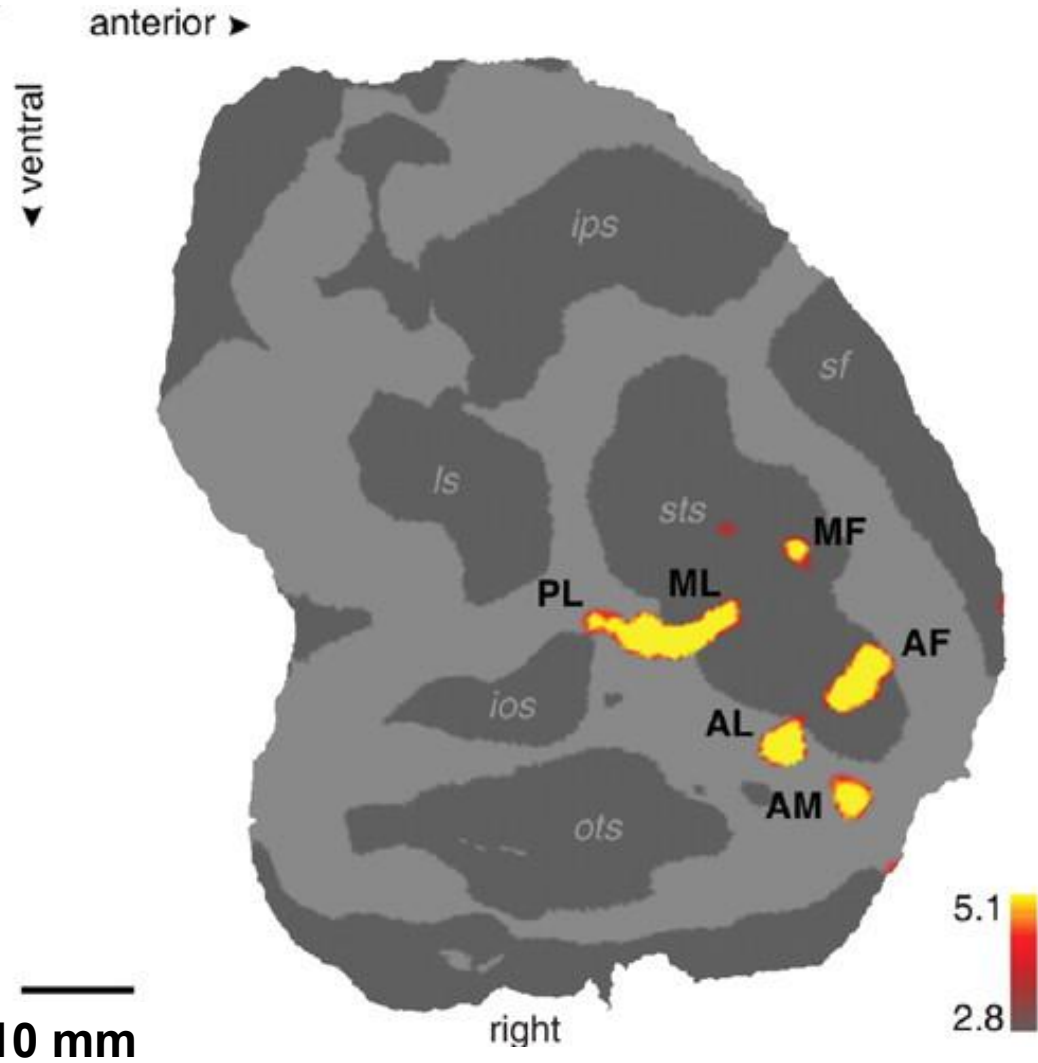
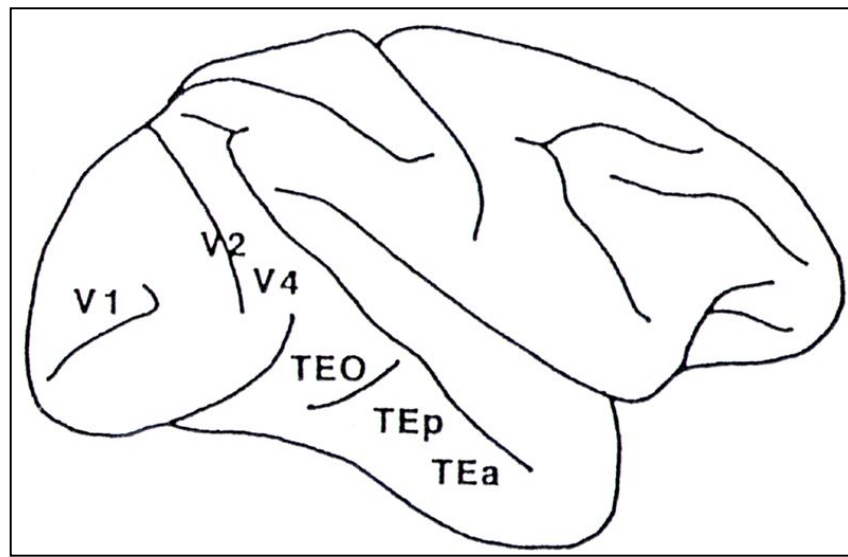


わかったことは

1. 隣接する細胞は異なる物体選択性を持っているように見えるが、一つのコラムの中にある細胞は共通の反応特性を持っている.

2. 共通の反応特性はコラムごとに異なっている.

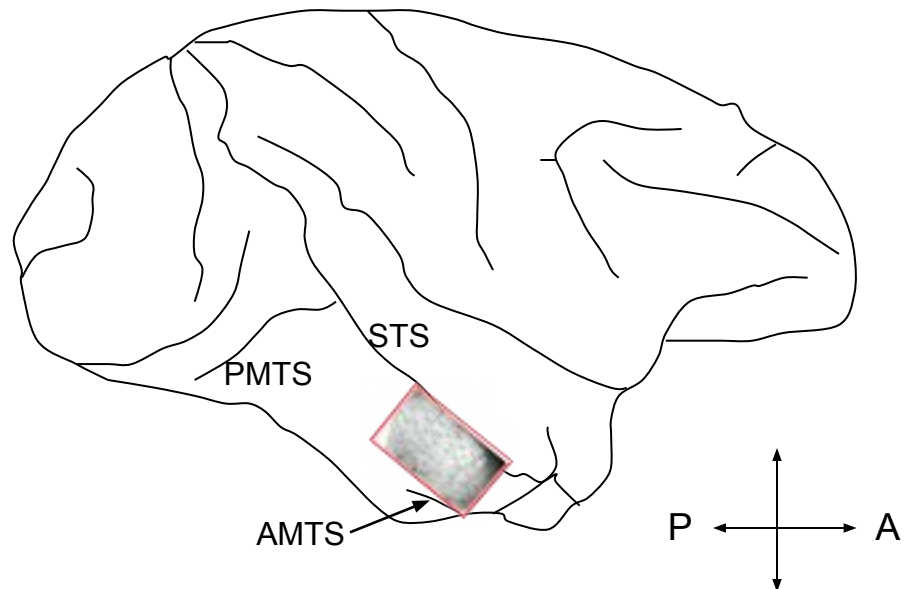
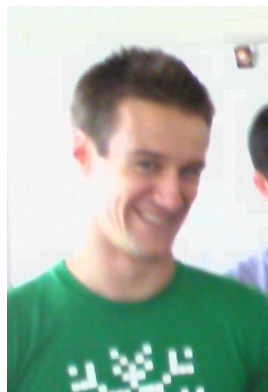




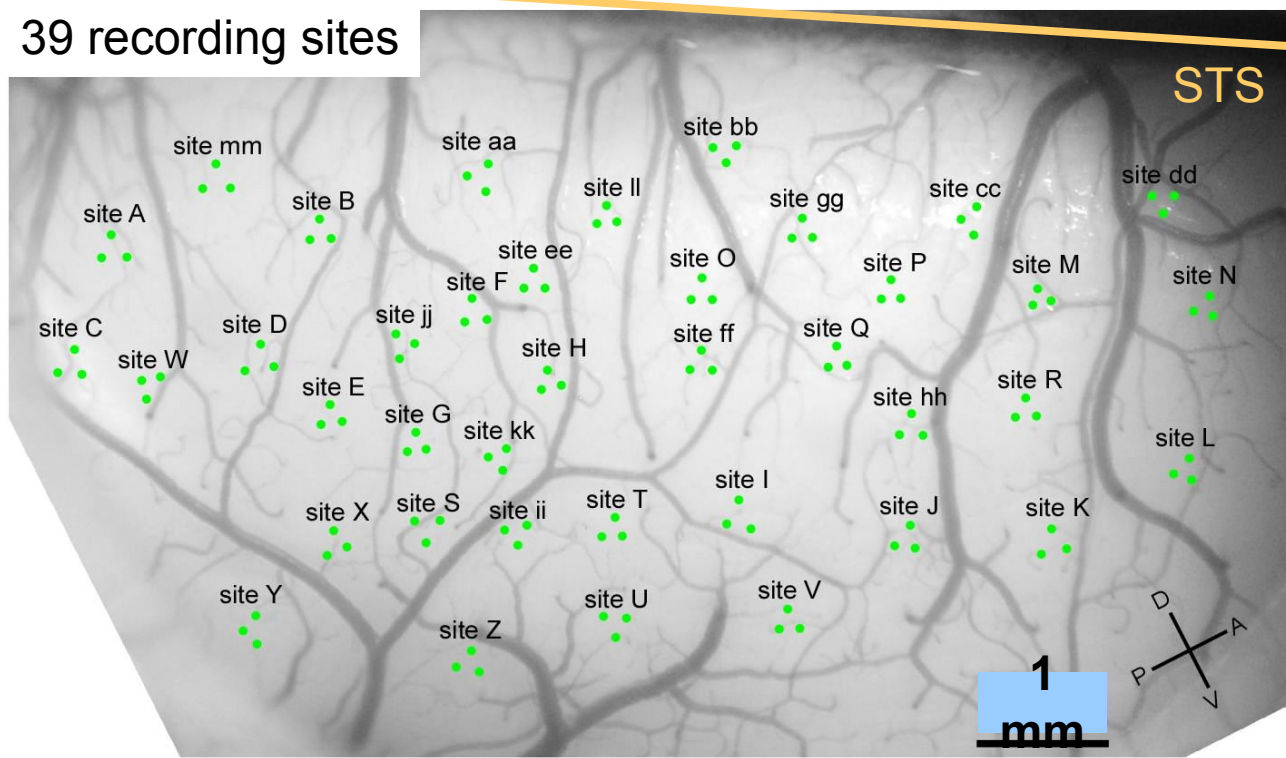
Takayuki Sato



Mark Lescroart



39 recording sites



Stimulus set

face
s

Scrambled
faces

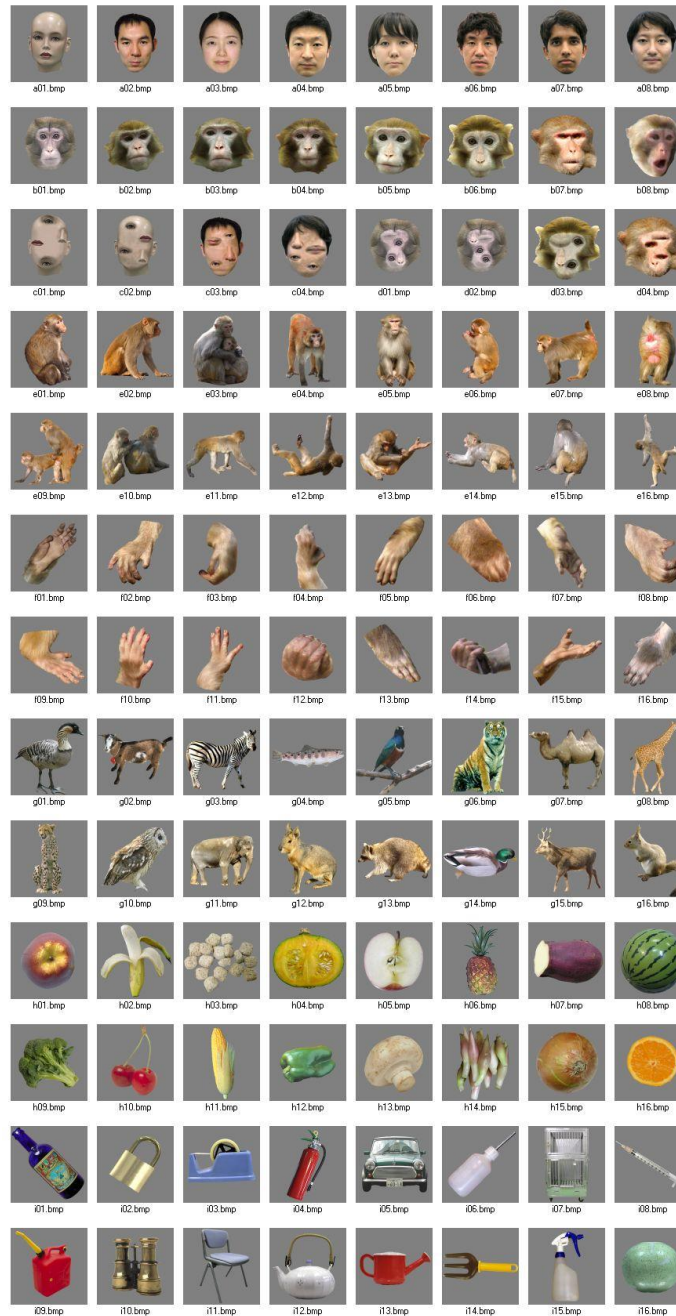
Monkey
bodies

Monkey
hands

Animal
bodies

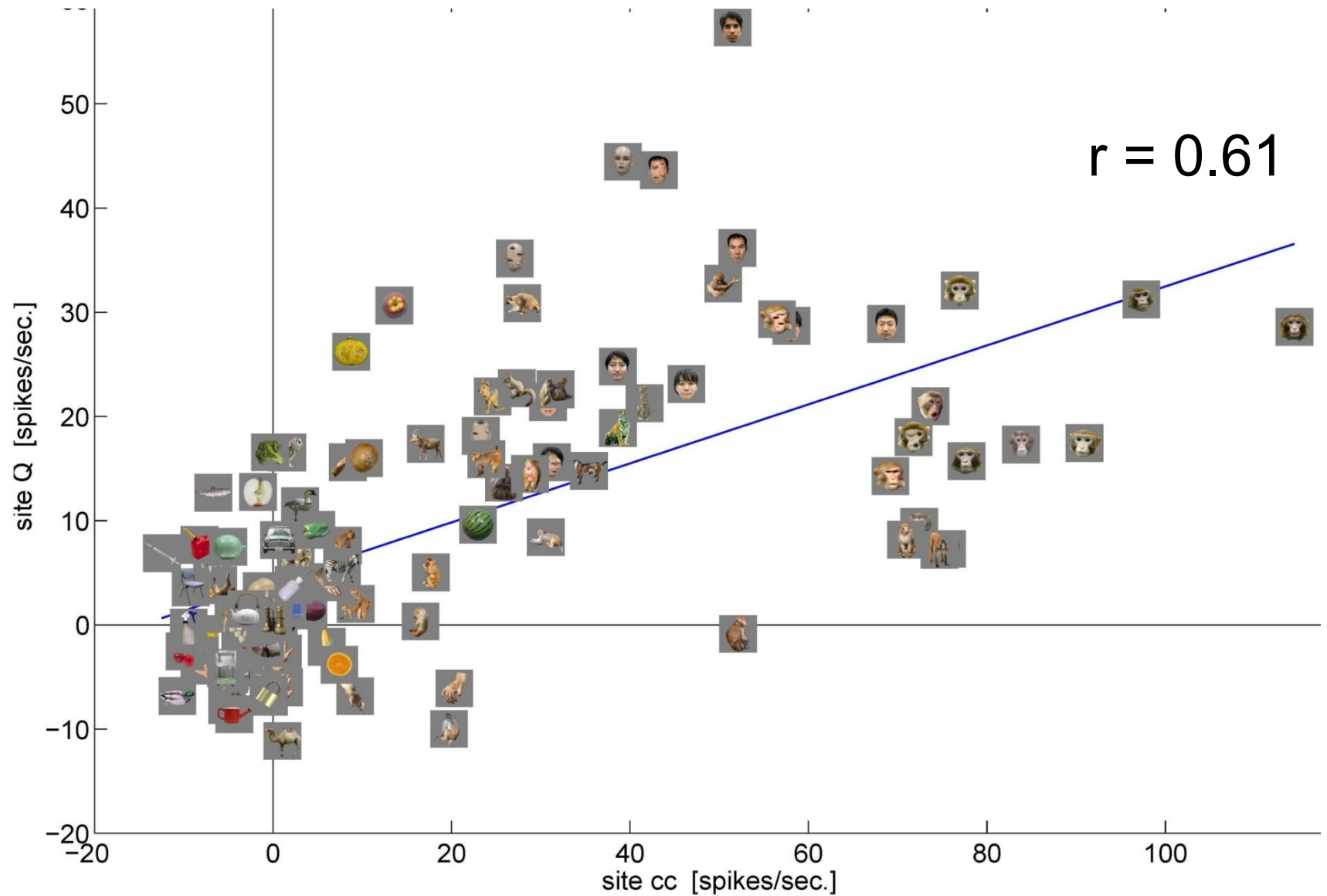
Food
s

Artificial
objects

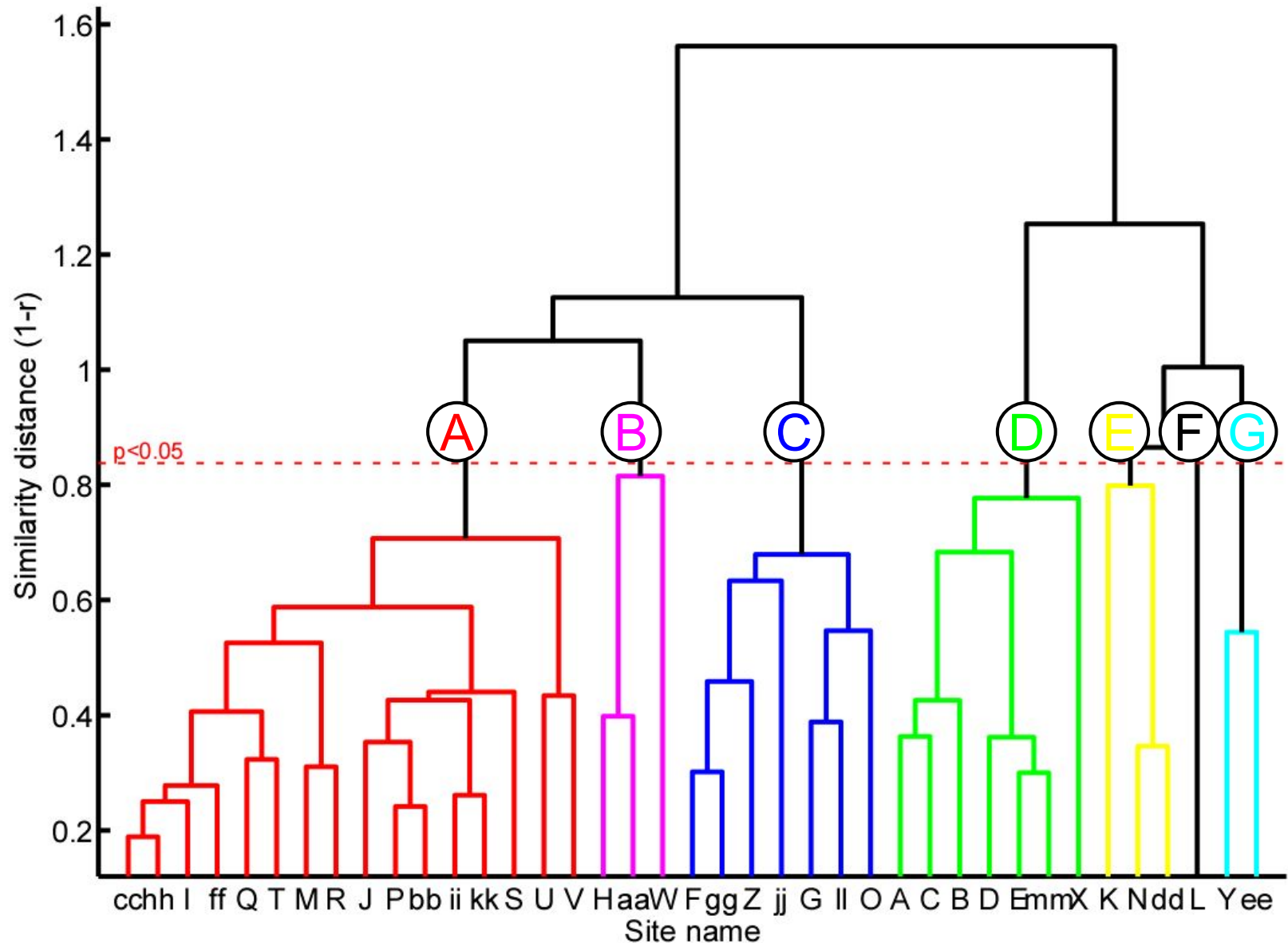


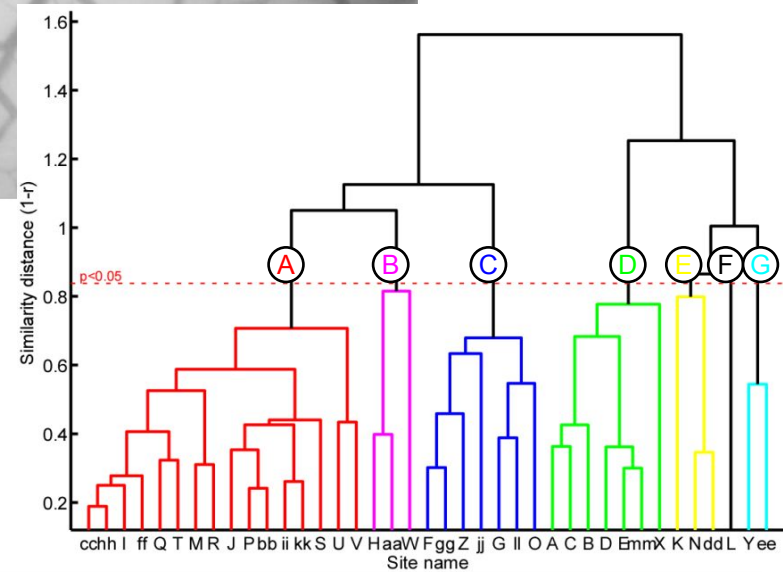
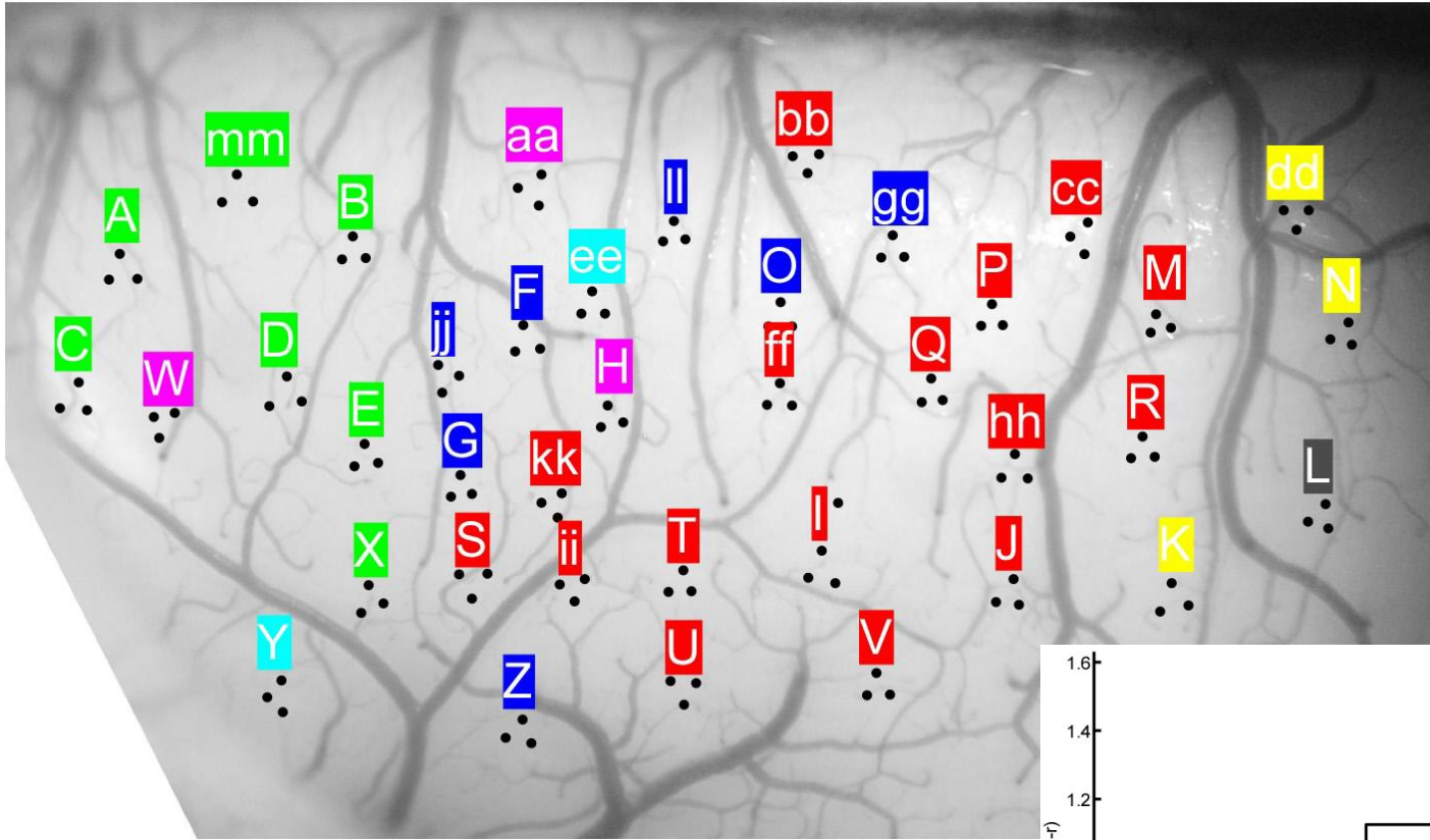
10
degree

Correlation coefficient as a similarity measure

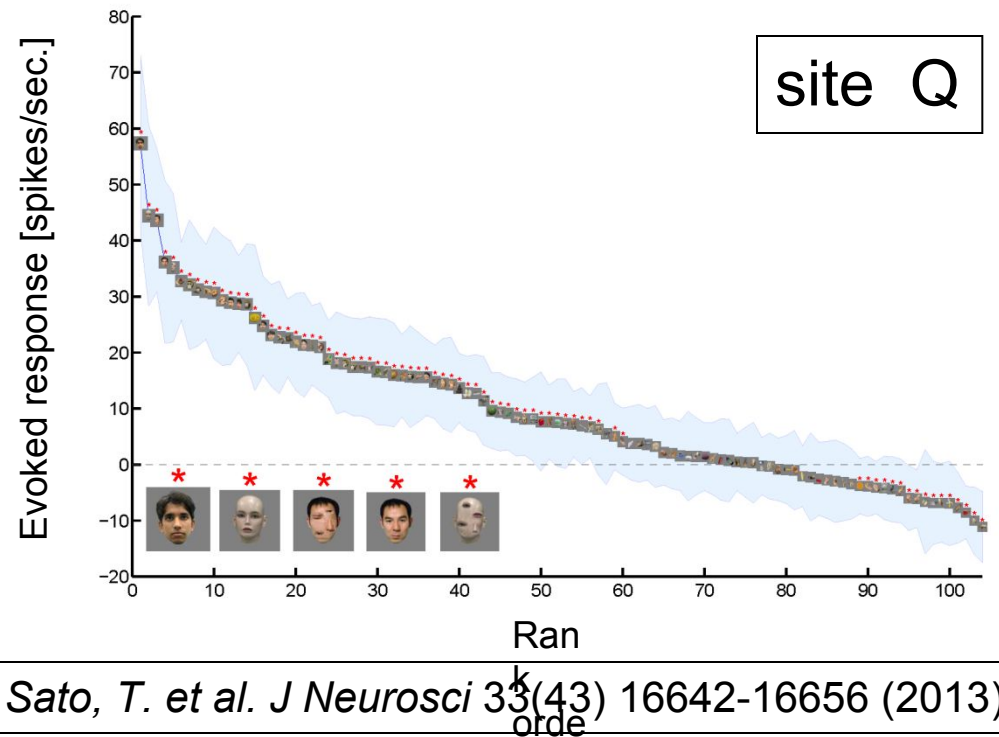
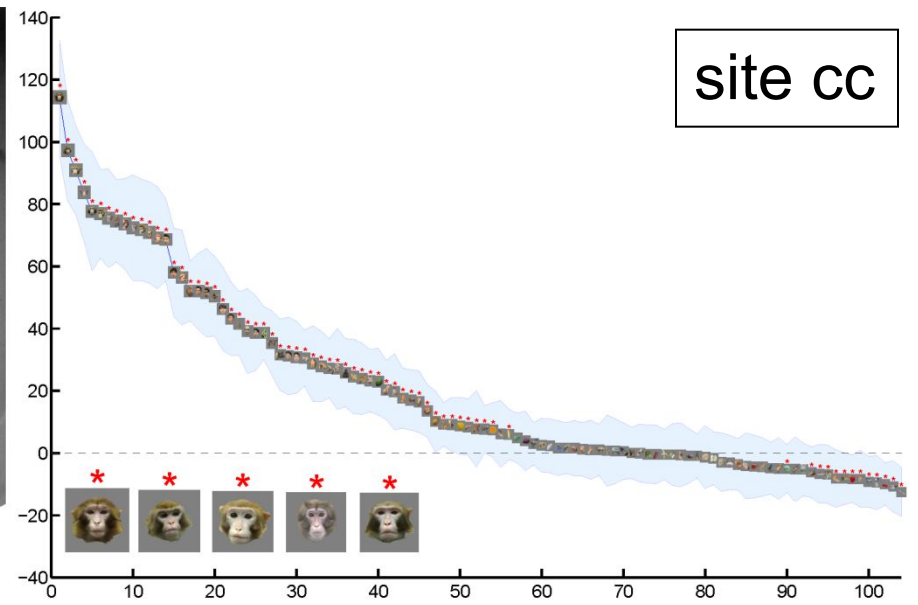
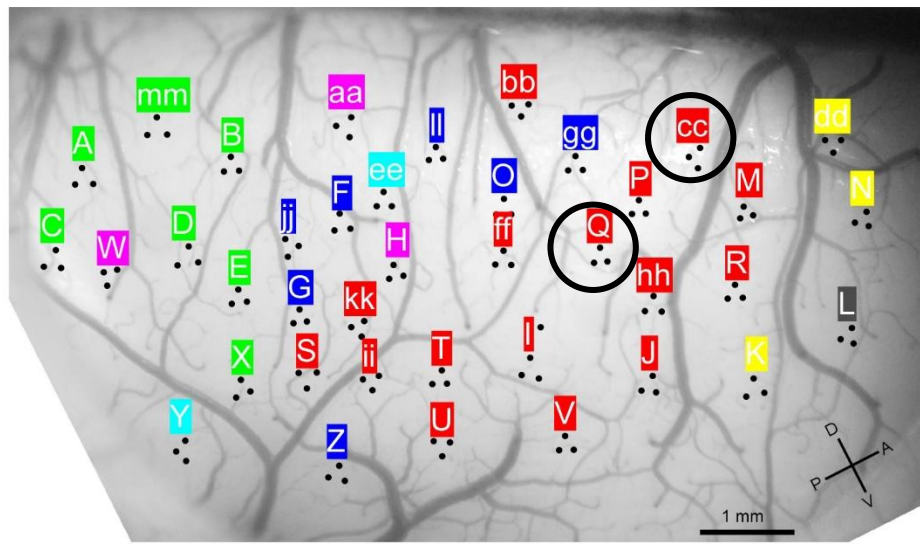


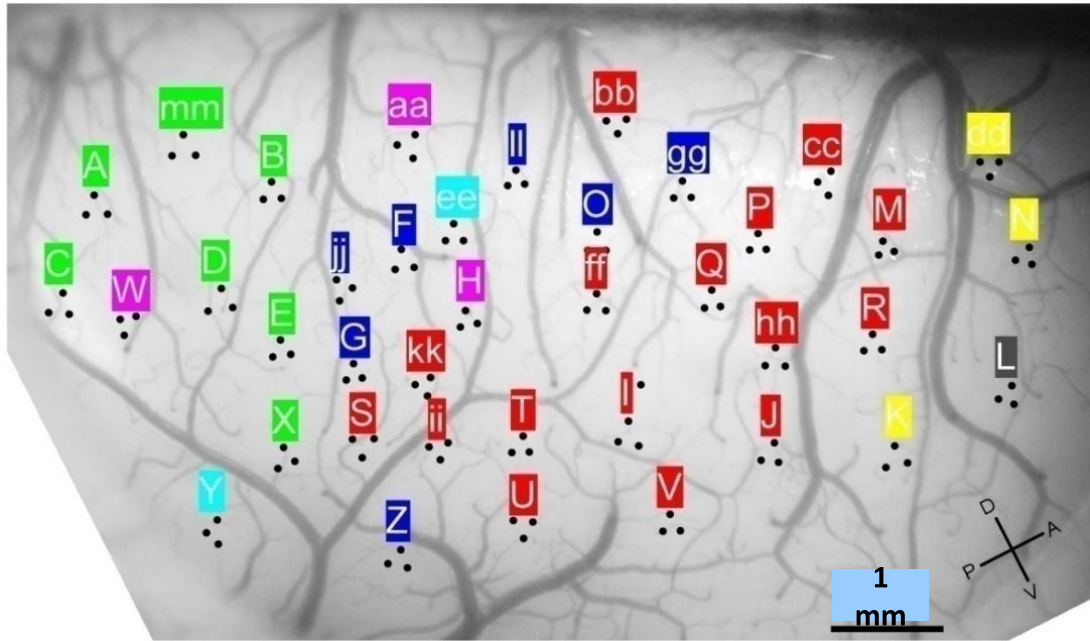
Recording sites are classified into 7 clusters.



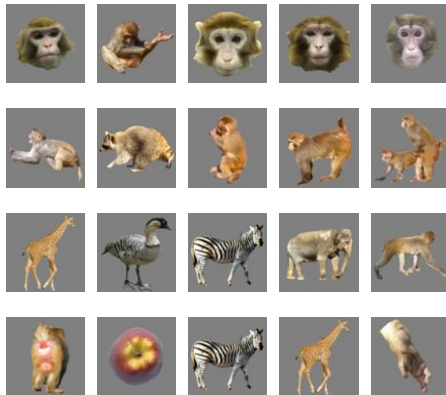


Sato, T. et al. *J Neurosci* 33(43) 16642-16656 (2013)

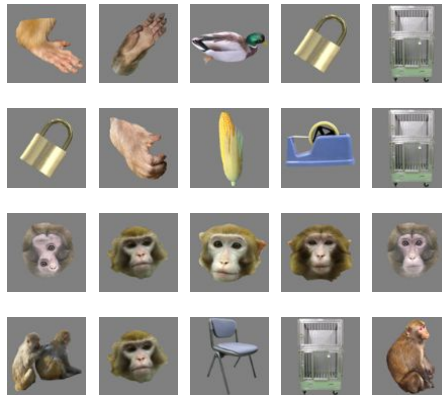




Best 5



Worst 5

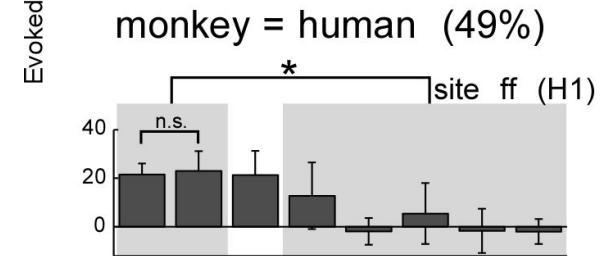
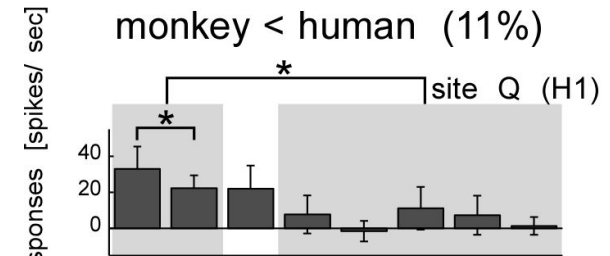
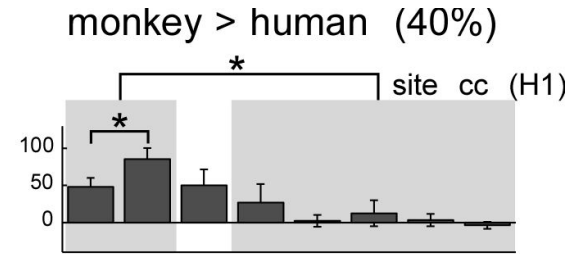


Domain A

Domain C

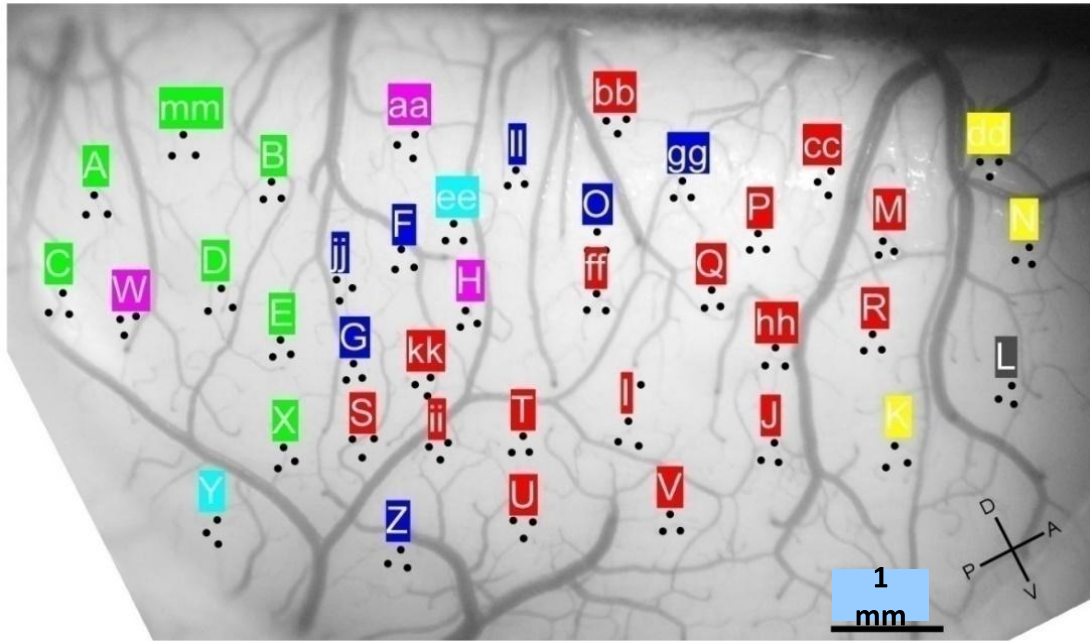
Domain D

Domain E

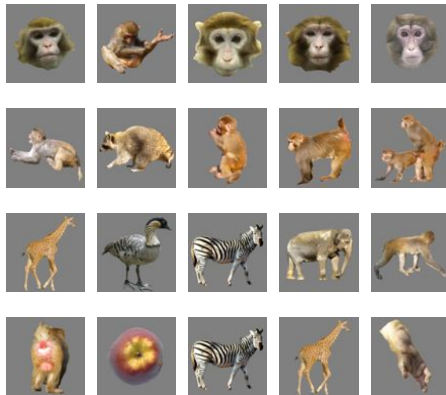


まとめ

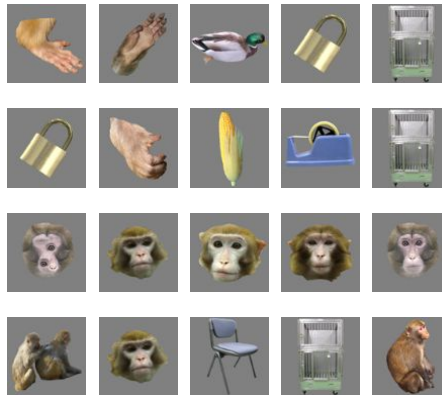
1. 単一細胞には固有の反応特性とコラムの中の細胞で共有する特性を持つ
2. 顔のような物体カテゴリに関係する視覚特徴に関係するコラムは皮質上に集まってドメインを作っている。
3. 物体表現は細胞、コラム、ドメインのように階層的である。



Best 5



Worst 5

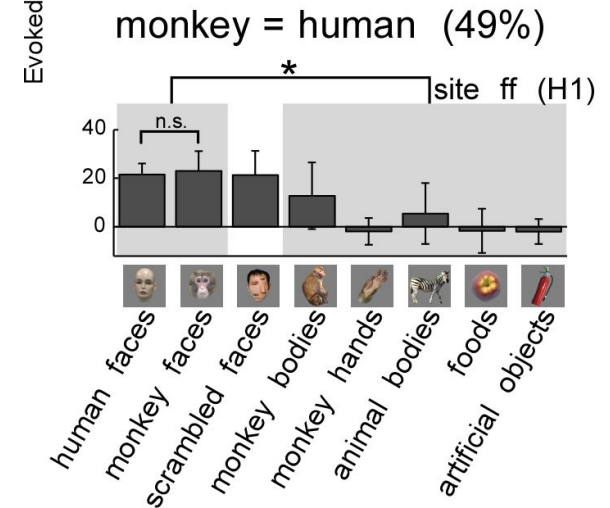
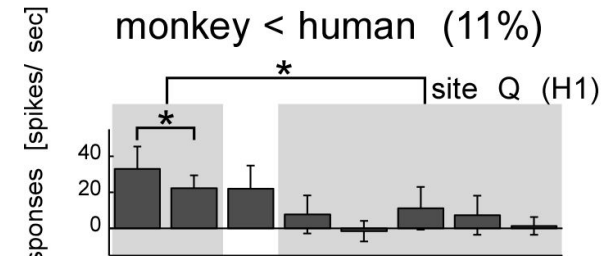
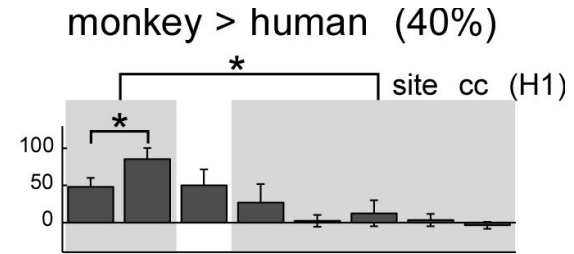


Domain A

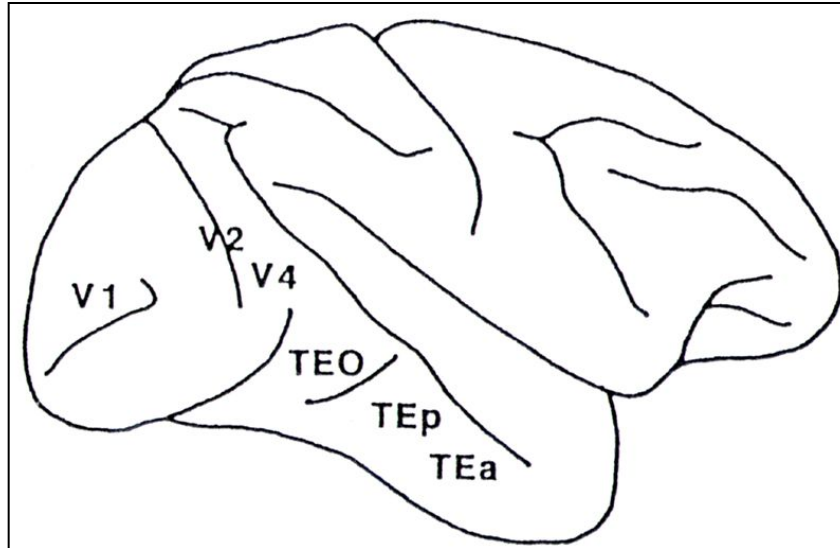
Domain C

Domain D

Domain E



サル高次視覚野の機能と構造



1. 高次視覚野の機能構造
2. 機能単位が表現する視覚特徴
3. 物体表現のダイナミクス

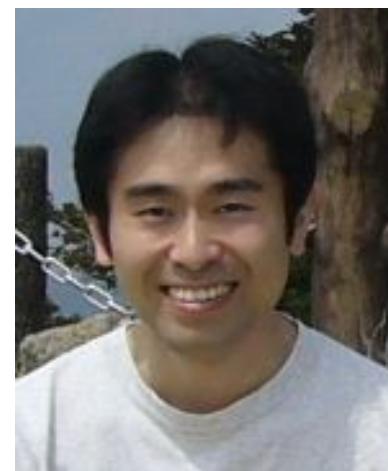
個々のコラムはどのような視覚特徴を符号化しているか？



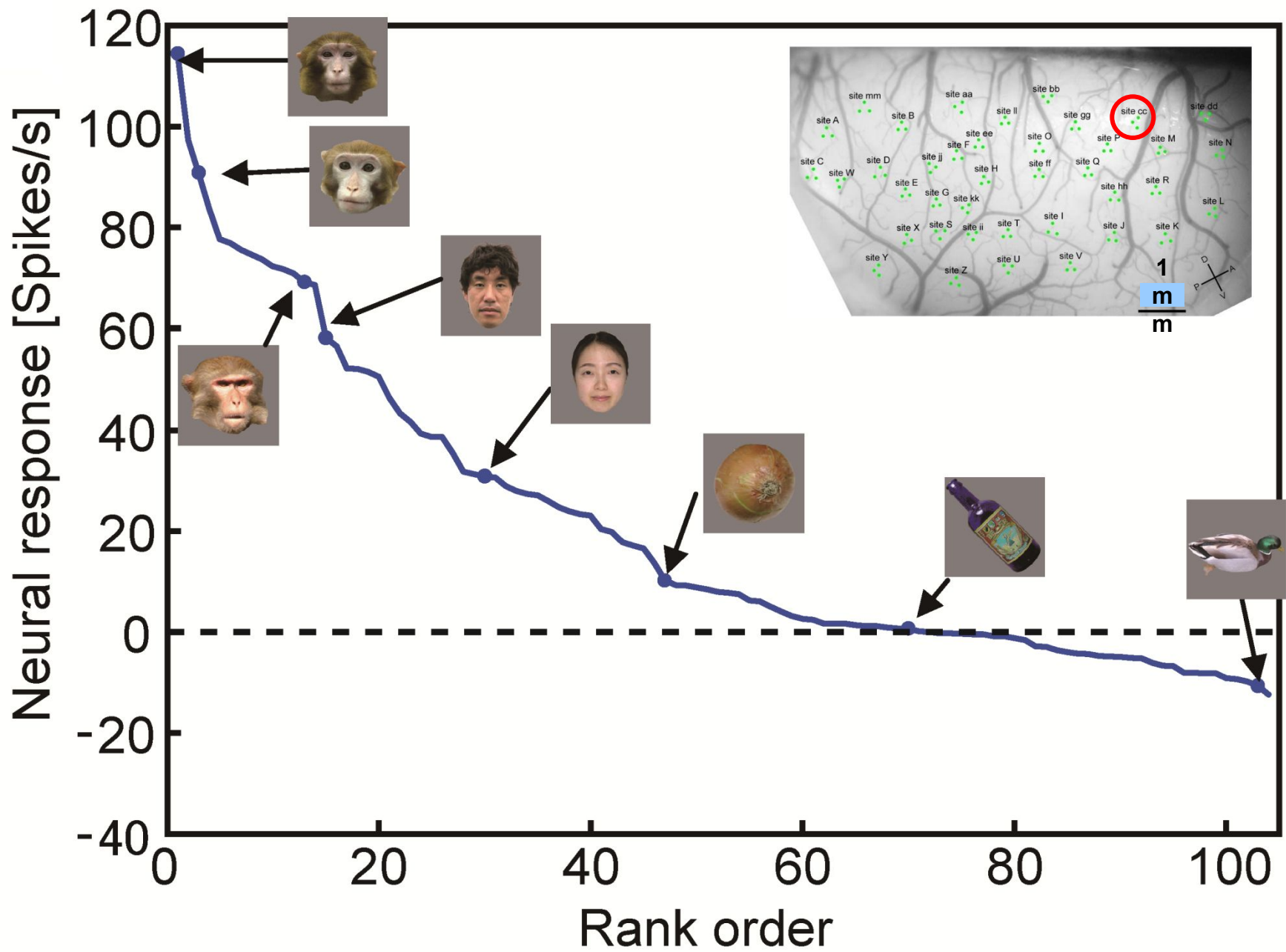
Shimon Ullman

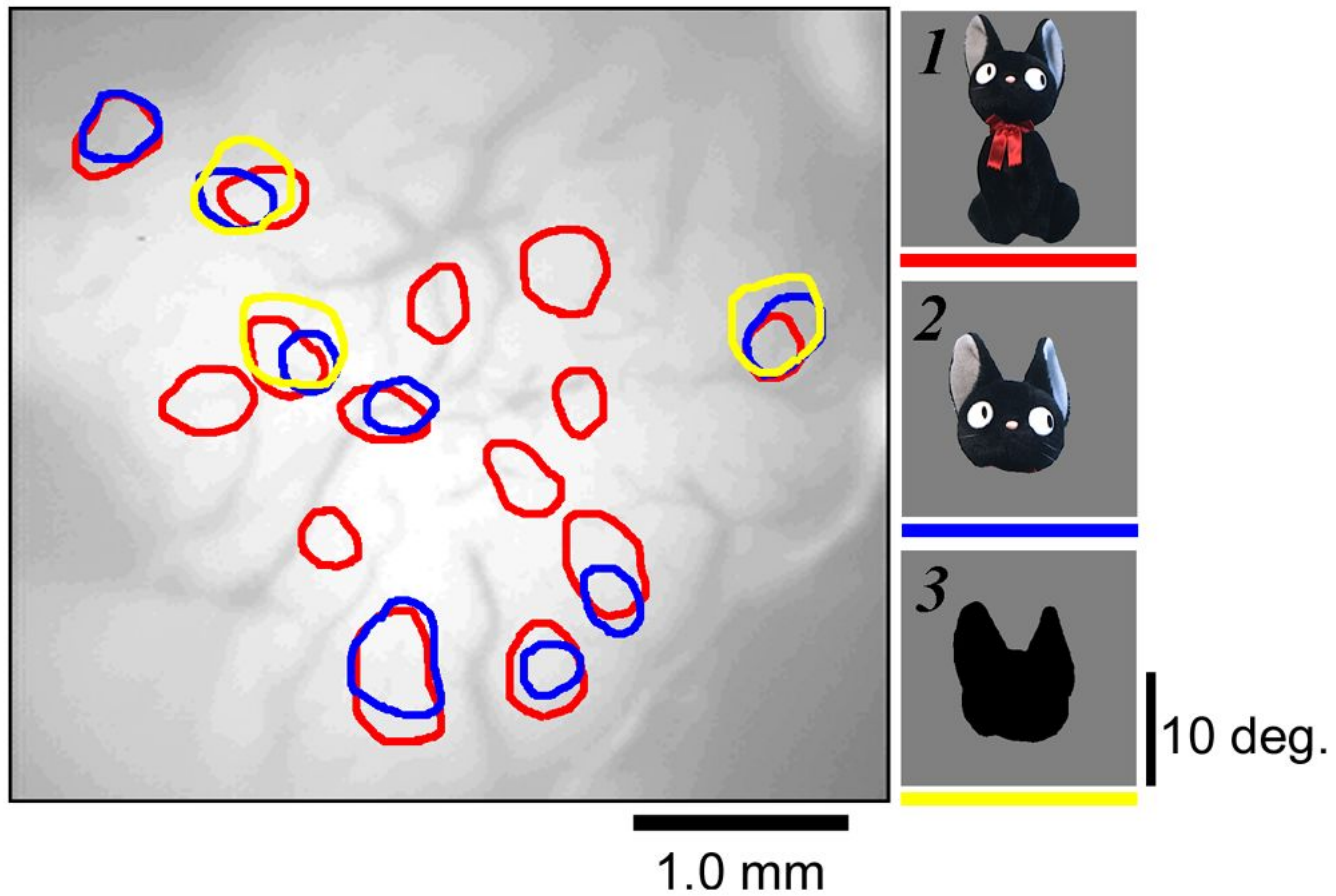


Michel Vidal-Naquet



Takashi Owaki

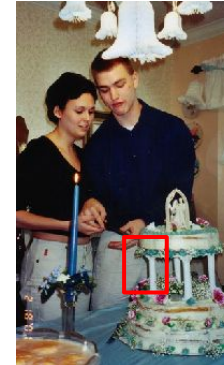
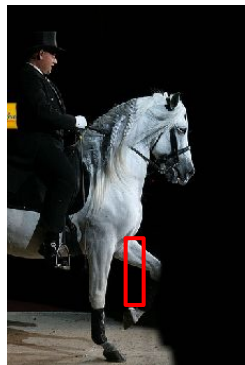
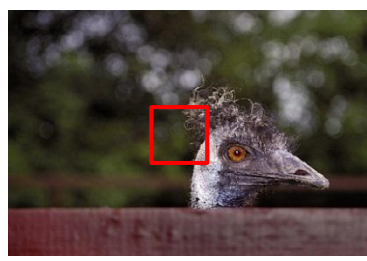
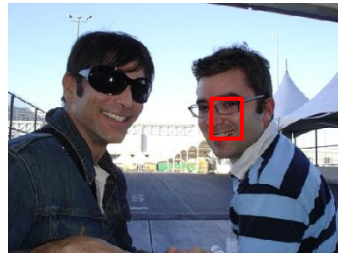
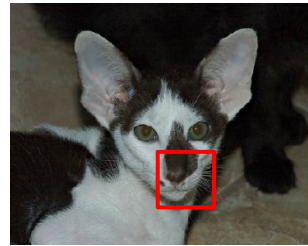
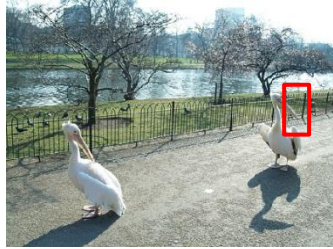




仮定

コラムが符号化している視覚特徴に近いものは、自然画像のどこかに見つかるはずだ。

560,000 candidate fragments extracted from 7753 natural images



この方法の詳細は未発表データのため公表は差し控えますが、自然画像の断片の中から、神経応答の選択性をよく説明するものを見つけ出すことに成功しました。これらの画像断片は、局所的な方位と色の組み合わせで構成されていました。

まとめ

1. 単一細胞には固有の反応特性とコラムの中の細胞で共有する特性を持つ
2. 顔のような物体カテゴリに関係する視覚特徴に関係するコラムは皮質上に集まってドメインを作っている。
3. 物体表現は細胞、コラム、ドメインのように階層的である。
4. コラムで表現されている視覚特徴は局所的な方位と色の組み合わせ

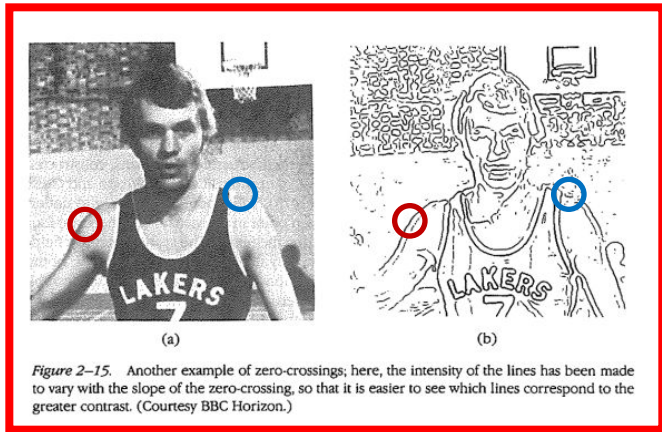
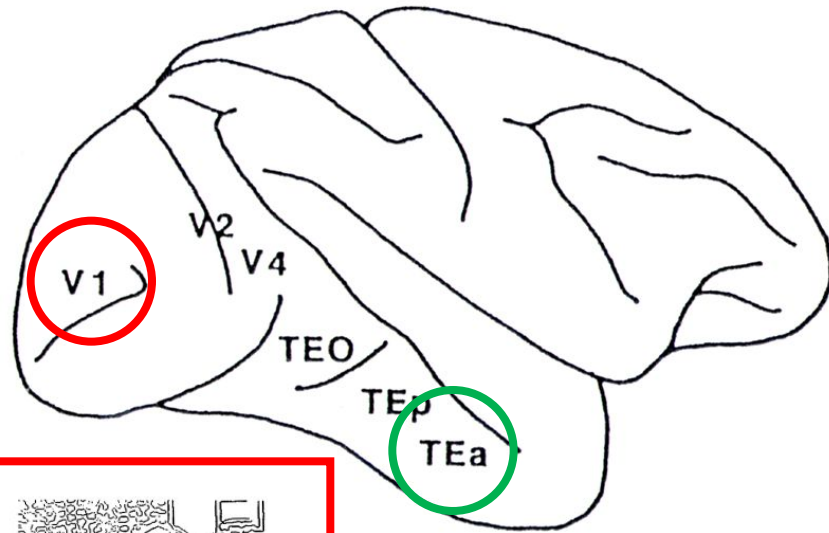
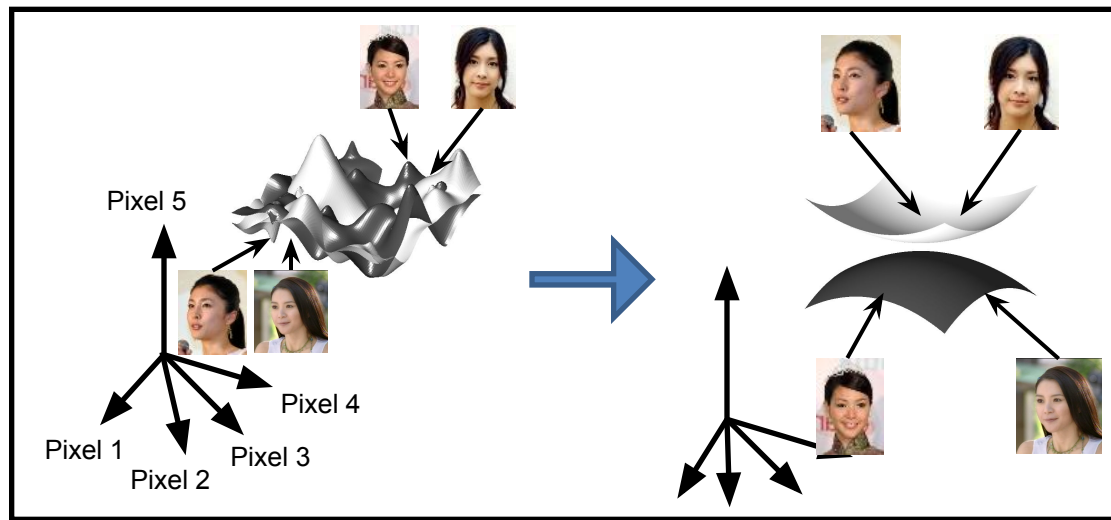
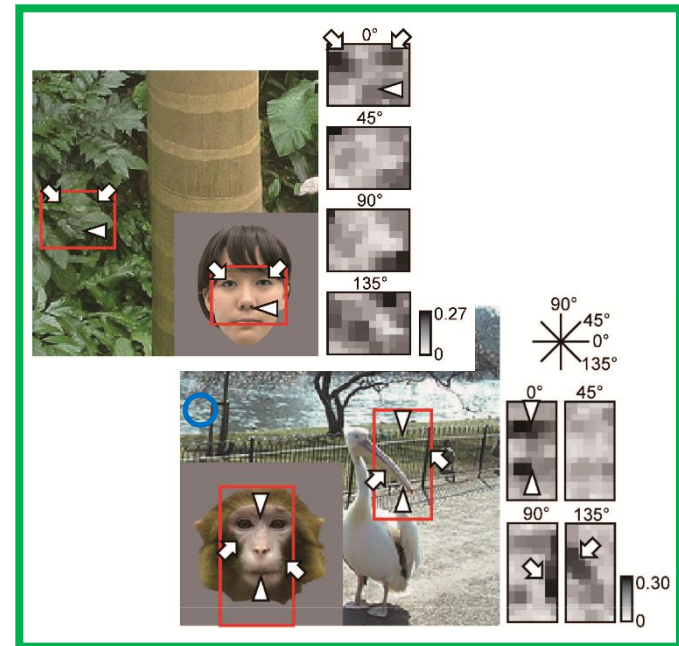


Figure 2-15. Another example of zero-crossings; here, the intensity of the lines has been made to vary with the slope of the zero-crossing, so that it is easier to see which lines correspond to the greater contrast. (Courtesy BBC Horizon.)



ImageNet Classification with Deep Convolutional Neural Networks

DCNN

Alex Krizhevsky

University of Toronto

kriz@cs.utoronto.ca

Ilya Sutskever

University of Toronto

ilya@cs.utoronto.ca

Geoffrey E. Hinton

University of Toronto

hinton@cs.utoronto.ca

Abstract

We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called “dropout” that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.



mite

container ship

motor scooter

leopard

	mite
	black widow
	cockroach
	tick
	starfish

	container ship
	lifeboat
	amphibian
	fireboat
	drilling platform

	motor scooter
	go-kart
	moped
	bumper car
	golfcart

	leopard
	jaguar
	cheetah
	snow leopard
	Egyptian cat



grille

mushroom

cherry

Madagascar cat

	convertible
	grille
	pickup
	beach wagon
	fire engine

	agaric
	mushroom
	jelly fungus
	gill fungus
	dead-man's-fingers

	dalmatian
	grape
	elderberry
	ffordshire bullterrier
	currant

	squirrel monkey
	spider monkey
	titi
	indri
	howler monkey

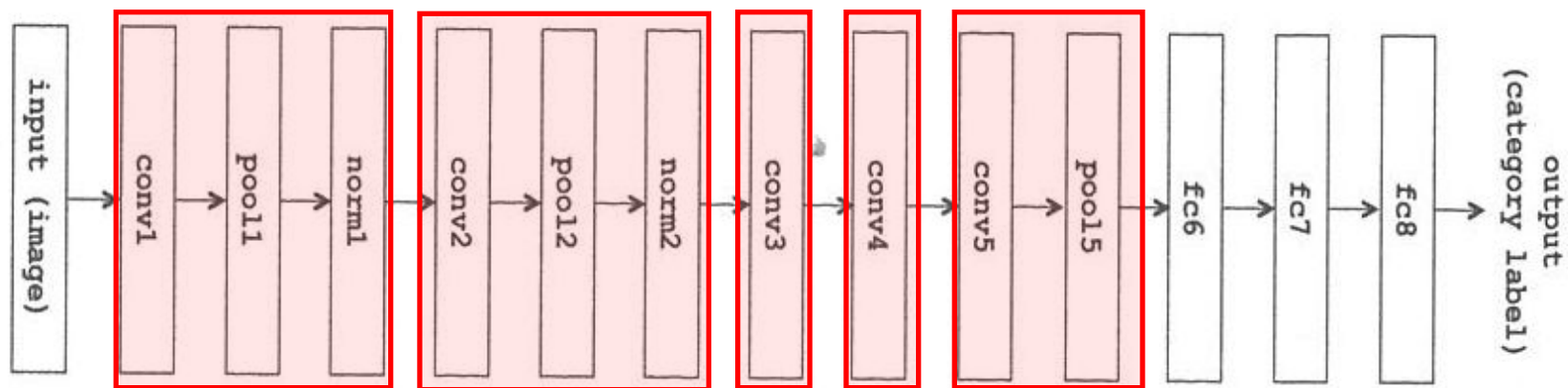
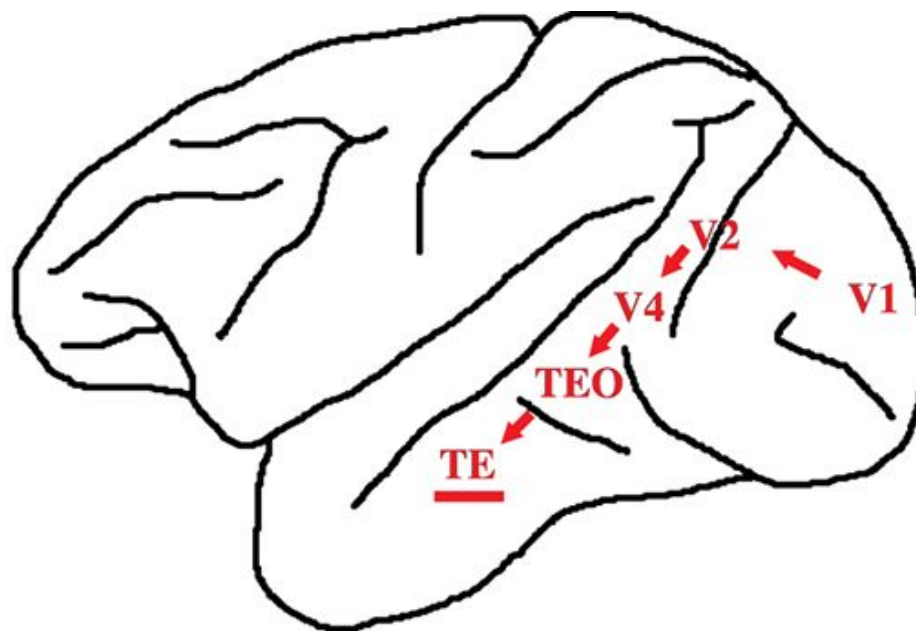
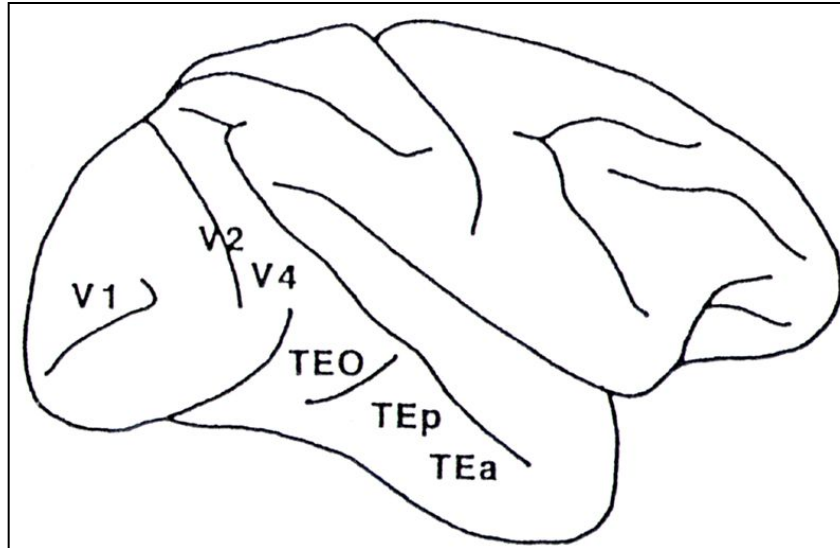


図 6.11 ILSVRC2012 での Krizhevsky らの畳込みネットの構造。表 6.2 も参照。

(岡谷貴之, “深層学習”)



サル高次視覚野の機能と構造



1. 高次視覚野の機能構造
2. 機能単位が表現する視覚特徴
3. 物体表現のダイナミクス

コラム内応答の受容野は動的に変化するか？

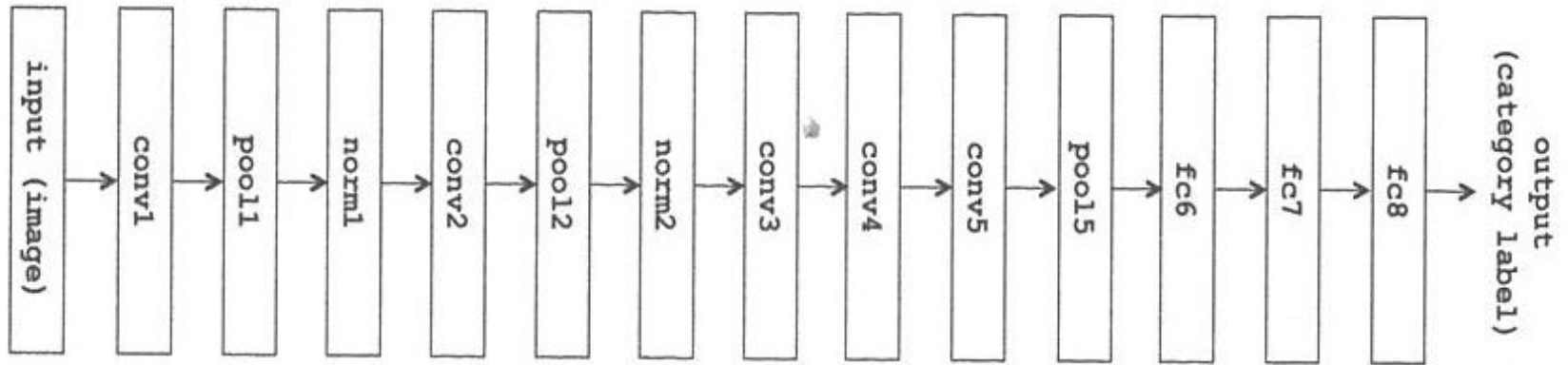
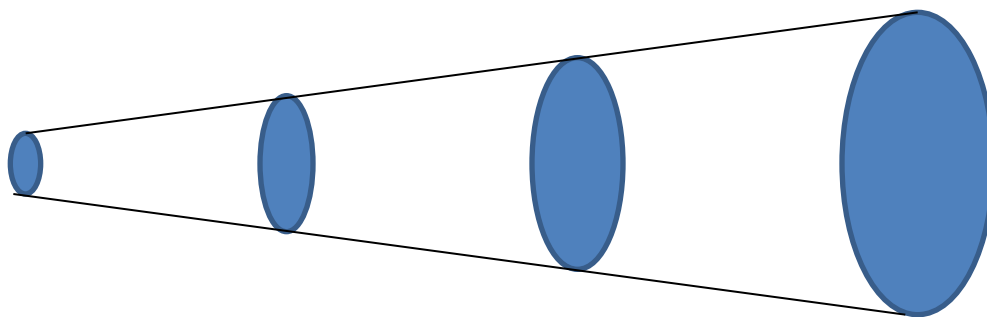
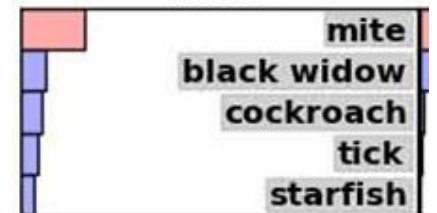


図 6.11 ILSVRC2012 での Krizhevsky らの畳込みネットの構造。表 6.2 も参照。

(岡谷貴之, “深層学習”)



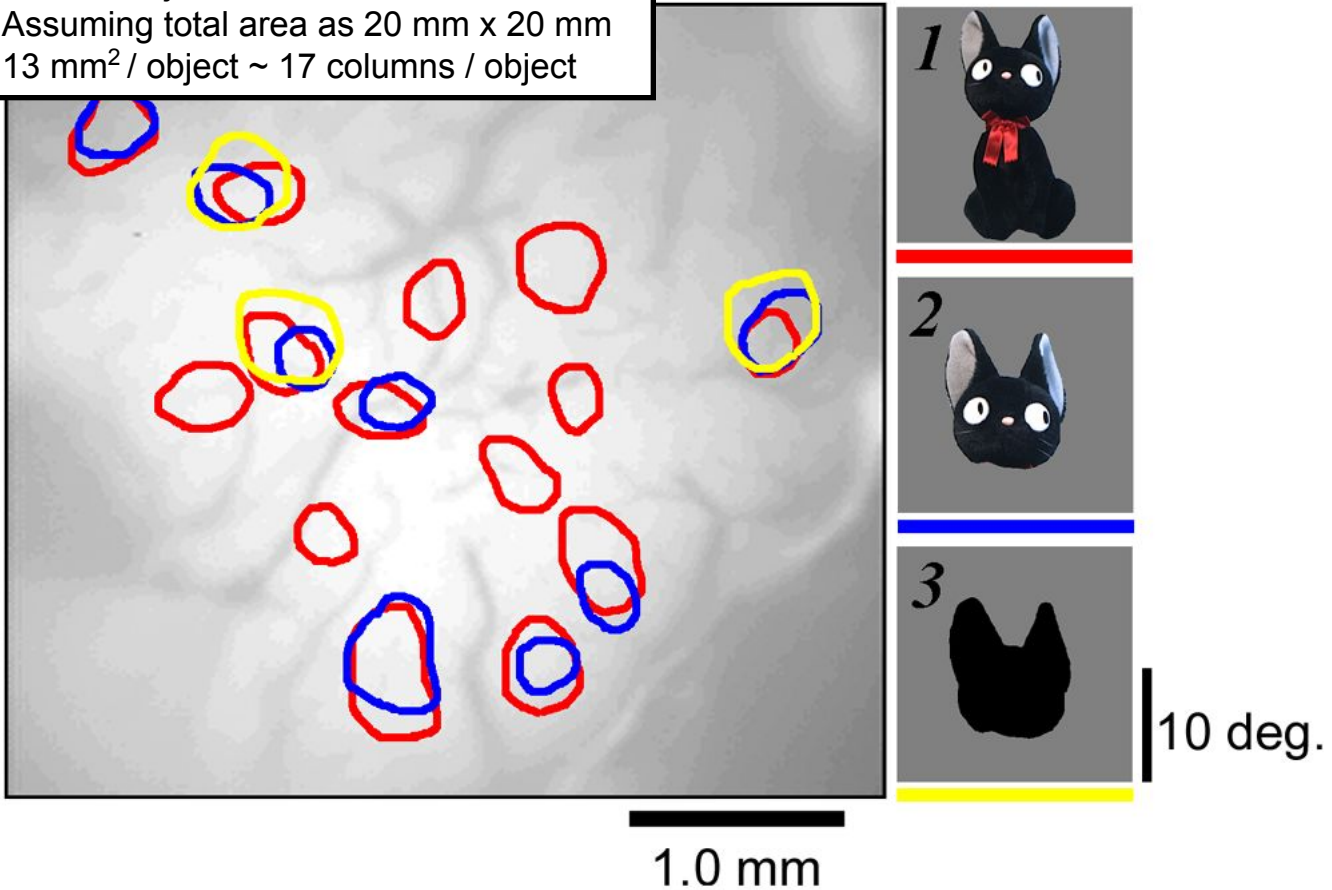
mite





ここで紹介したDeep convolutional Neural Network (Alex net)は脳における視覚情報処理の過程と似ている部分もありますが、同じではありません。特に、Alex Netが静的なネットワークであるのに対し、脳における視覚情報処理はダイナミックなプロセスで、物体像の表現もダイナミックに変化します。私の研究室では、実際、注意によって、物体像に対する反応が変化することを見出しました。詳細は未発表データのため公表は差し控えます。

3.3 % / Object
Assuming total area as 20 mm x 20 mm
13 mm² / object ~ 17 columns / object



Tsunoda, et al. Nat. Neurosci. . 4, 832-838 (2001)

この前のスライドに示すように、脳における物体像の表現はスパースです。すなわち、高々 10^2 個程度の数のコラムを使って、個々の物体像が表現されている。AlexNetとの関係を考える上で、物体像の表現はスパースかどうか？もし、そうでないとすれば、なぜスパースにならないのか？一つのカギは情報表現が静的であることが関係しているかもしれません。例えば、「注意によって」ターゲットに目を向けることによって、位置による不変性をネットワークで実現する必要がなくなります。それが、表現のスパース性に関係しているかもしれません。

脳により近いネットワークは、注意の対象を視野の中で見つけるための分解能の粗いネットワークと視野の中心にある物体像を同定する分解能の高いネットワークの二つのネットワークが相互の結合したものかもしれない。