

論文紹介:

「皮質の自発活動が内部モデルの環境への最適化の証拠を明らかにする」

Berkes, P., Orban, G., Lengyel, M., and Fiser, J. (2011).
Spontaneous cortical activity reveals hallmarks
of an optimal internal model of the environment.
Science, 331(6013):83.87

(所内ゼミでの発表資料の修正版)

2012-06-01

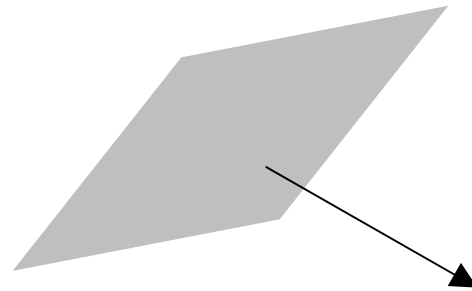
産業技術総合研究所 一杉裕志

ニューロンの自発活動がベイズモデルにおける事前分布と一致することなどを前提にして、フェレットのV1の活動が成長するにつれて外界に適応していく証拠を実験的に示した論文を紹介します。

概要：

脳は、感覚入力を解釈し行動の準備をするために、外部環境の内部モデルを維持している。行動レベルの研究では脳の内部モデルは環境の統計に最適化されることが示されているが、そのニューロンによる実現方法は知られていない。我々は感覚野での情報処理のベイズモデルに基づいて、刺激で誘発される活動を内部モデルによる推論結果に、自発活動を活動の事前分布に関連付けた。そして、もし内部モデルが統計的に最適ならばこれらが一致するはずだと予測した。この予測を検証するため、発達段階における覚醒したフェレットの視覚皮質の活動を分析した。自発活動と自然刺激による誘発活動との類似性は、年齢とともに増大した。このことは、内部モデルが自然刺激の統計に徐々に適応することを、ニューロンレベルで示している。

先行研究：脳がベイズ統計に従っている証拠の一例 [Weiss et al. 2002]



ひし形の運動方向の
認知に関わる錯視のモデル

Y. Weiss, E. P. Simoncelli, E. H. Adelson,
Motion illusions as optimal percepts. Nat. Neurosci.5, 598
(2002).

doi:10.1038/nn0602-858 pmid:12021763

デモ：「Moving Rhombus Displays by Yair Weiss」

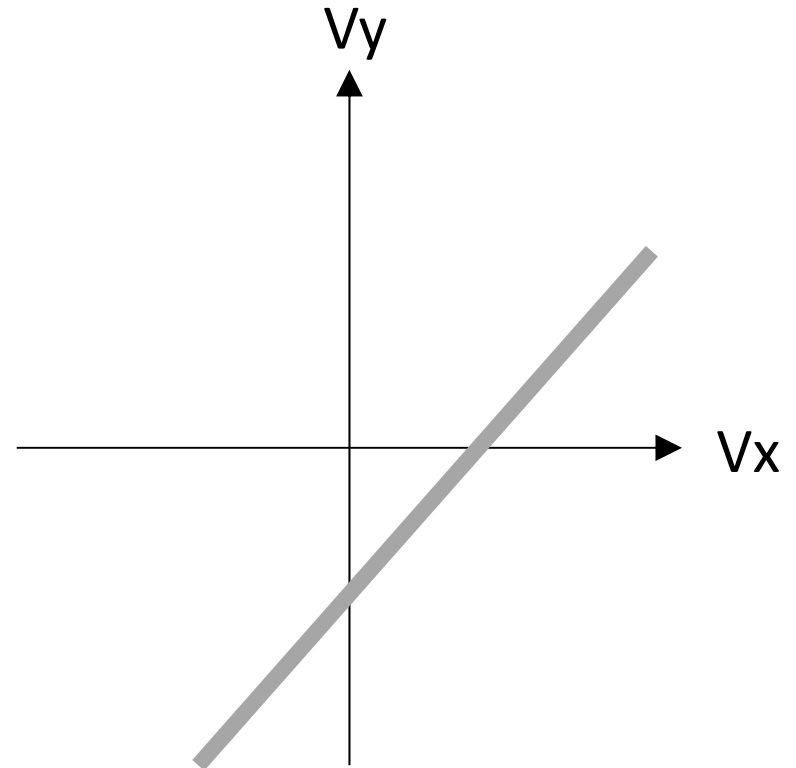
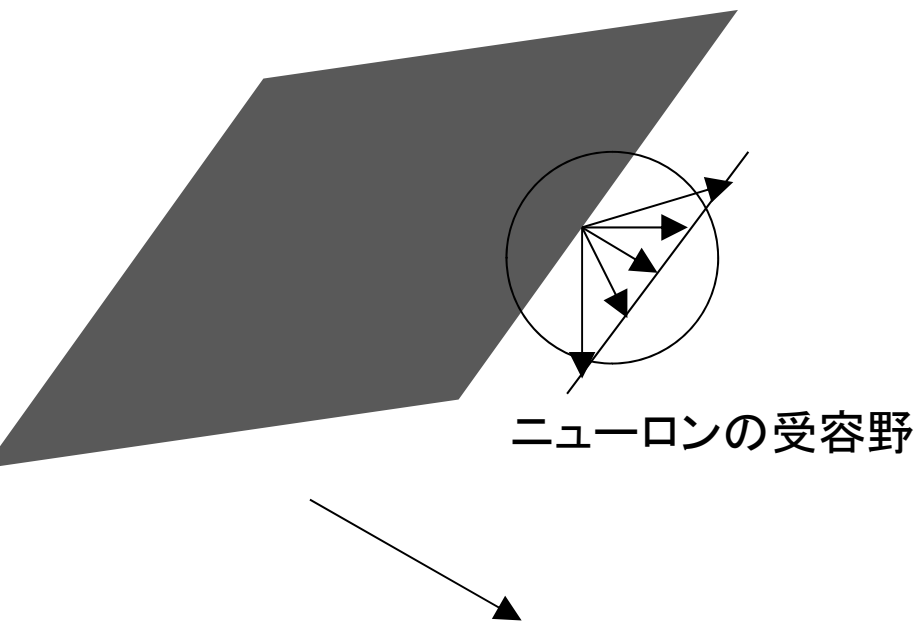
<http://www.cs.huji.ac.il/~yweiss/Rhombus/rhombus.html>

参考：日本語の解説

柴田和久, 神谷之康 ASCONE 講義-ベイズで読み解く知覚世界,
日本神経回路学会誌, 2007, 14(4):313-8

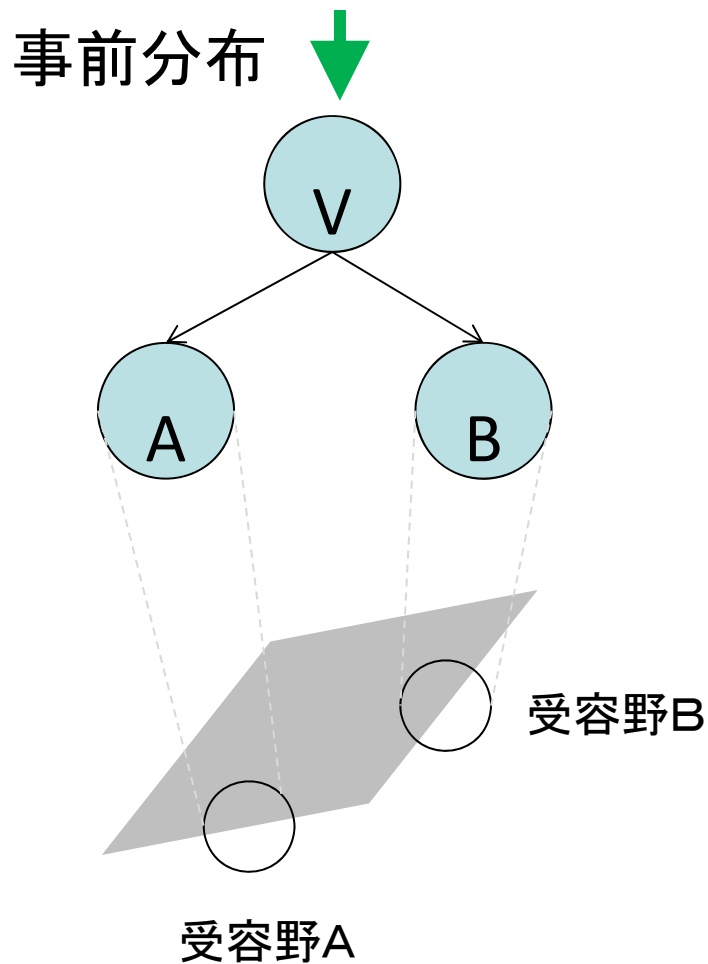
https://www.jstage.jst.go.jp/browse/jnns/14/4/_contents/-char/ja/

単一ニューロンからは物体全体の動きは分からない。



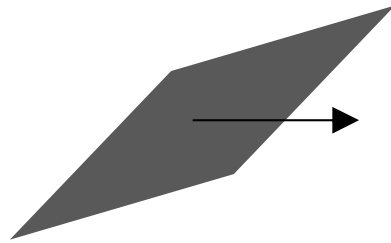
狭い受容野から推定される
物体の速度の確率分布

物体の速度を推定するモデル

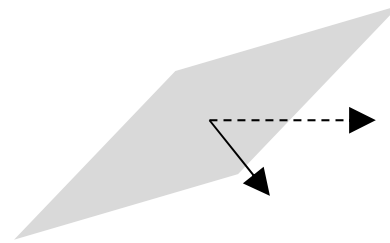


- 仮定:
 - 狭い受容野からの情報を統合して推定
 - コントラストが低いとノイズが大きい
 - 「物体の速度は小さい可能性が高い」という事前分布を持つ

コントラスト低(ノイズ大)だと 認識結果は事前分布に引きずられる。



本当の動き



コントラストが低いときは、
斜めに動いて見える。

この現象はベイズモデルによれば、
斜めに動いていると解釈した方が
動きが遅いので、
よりもっともらしいから、と説明できる。

今回紹介する論文の背景

- 行動のレベルでは、動物の振る舞いが「統計的最適」であることが多くの論文（[Weiss et al. 2002] など）で示されてきている。
- しかし、統計的最適さがニューロンでどう実現されているかは不明。
- 今回紹介する論文は、動物が成長するにつれてニューロン応答が「統計的最適」に近づいていく一例を示している。
 - （ニューロンでの実現方法を知る手掛かりになる。）

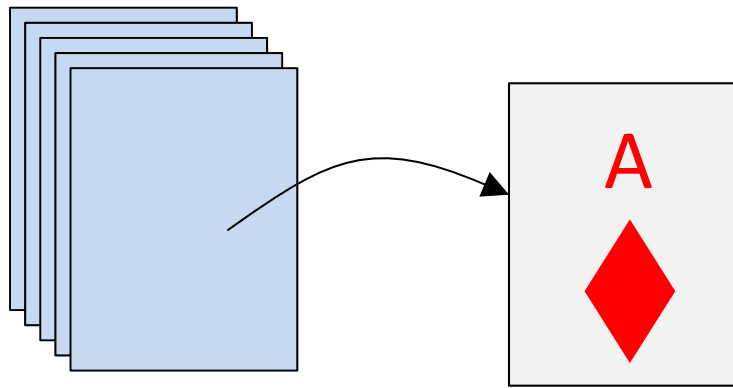
この論文が前提としていること

- 脳は「外部環境のモデル」を学習している。
- ニューロンの刺激への応答は、刺激のなんらかの「特徴」を表現している。
- 刺激に対するニューロン応答の頻度は、刺激に対する特徴の事後確率
 $P(\text{特徴} | \text{刺激、モデル})$
を表している。
- 自発発火は事前分布 $P(\text{特徴} | \text{モデル})$ を表している。

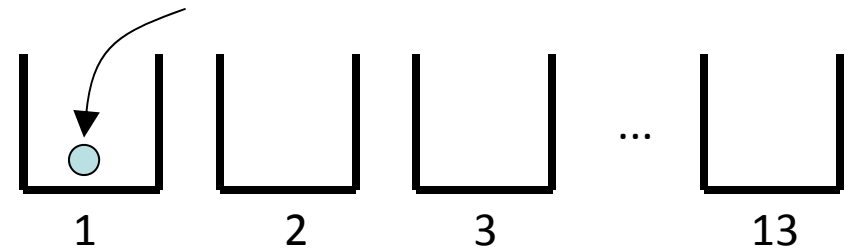
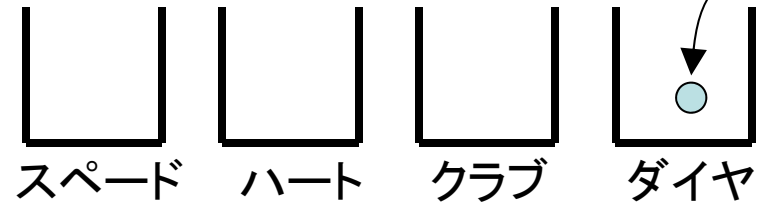
– [Weiss et al. 2002]などの実験や、いろいろな計算論的モデルが示唆している。

私なりに、わかりやすく例える
と・・・

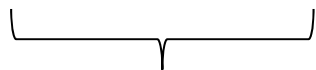
例えば、 脳は特徴の頻度を学習すると仮定



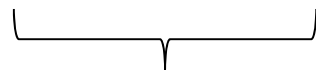
カードの山から1枚
引いては山に戻す。
山の中身は不明。



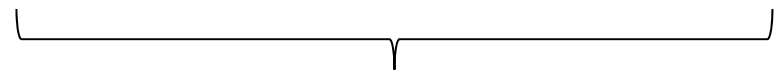
感覚器への入力 that 得られるたびに
特徴のヒストグラムを更新



外界

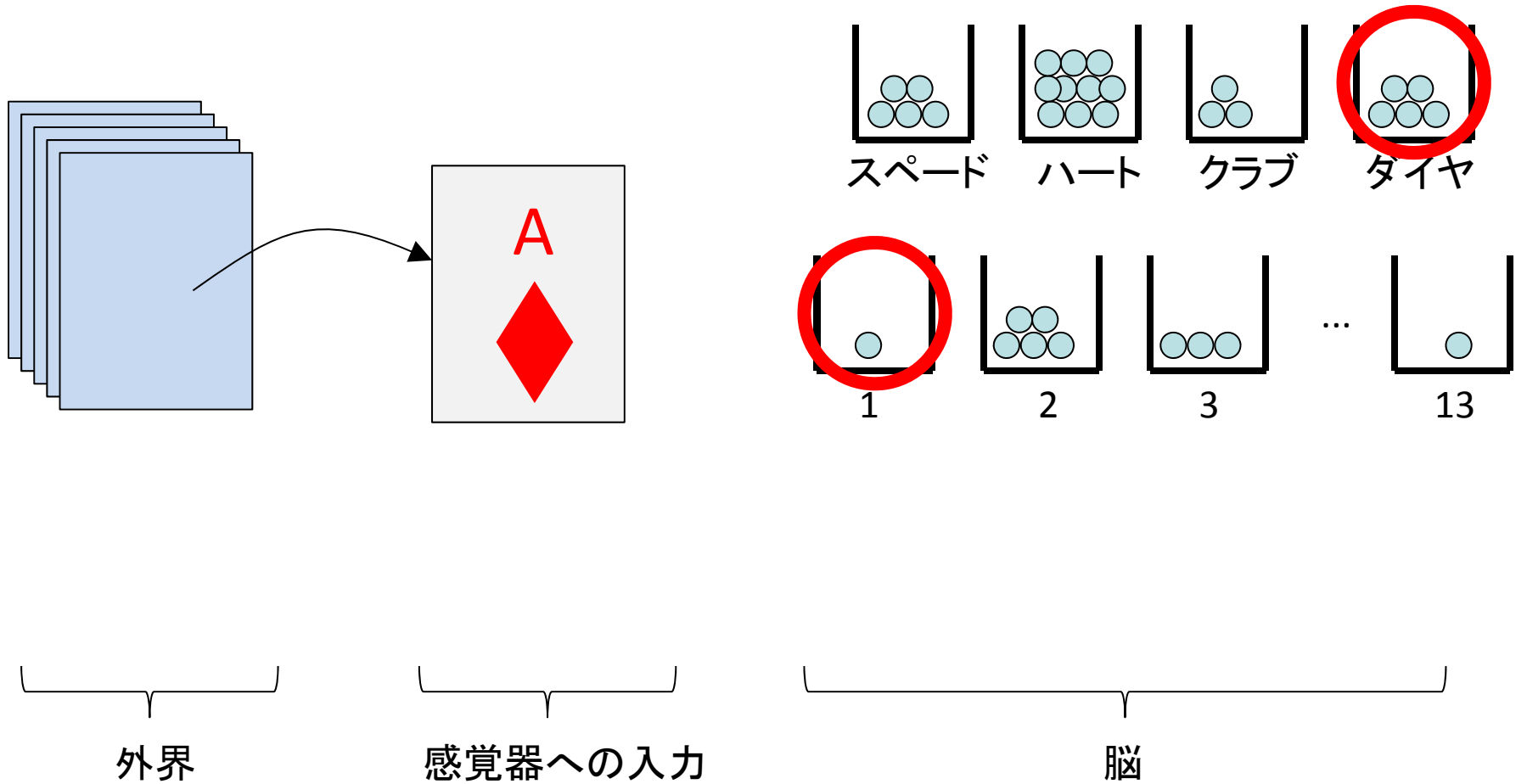


感覚器への入力

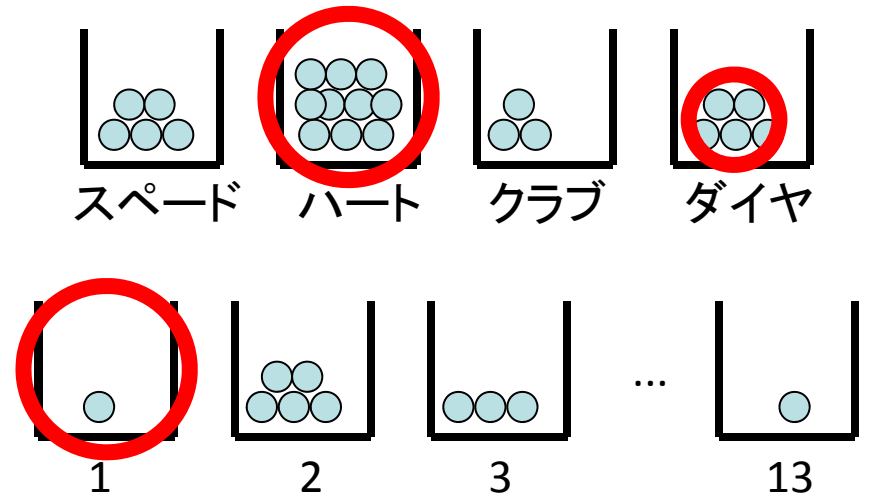
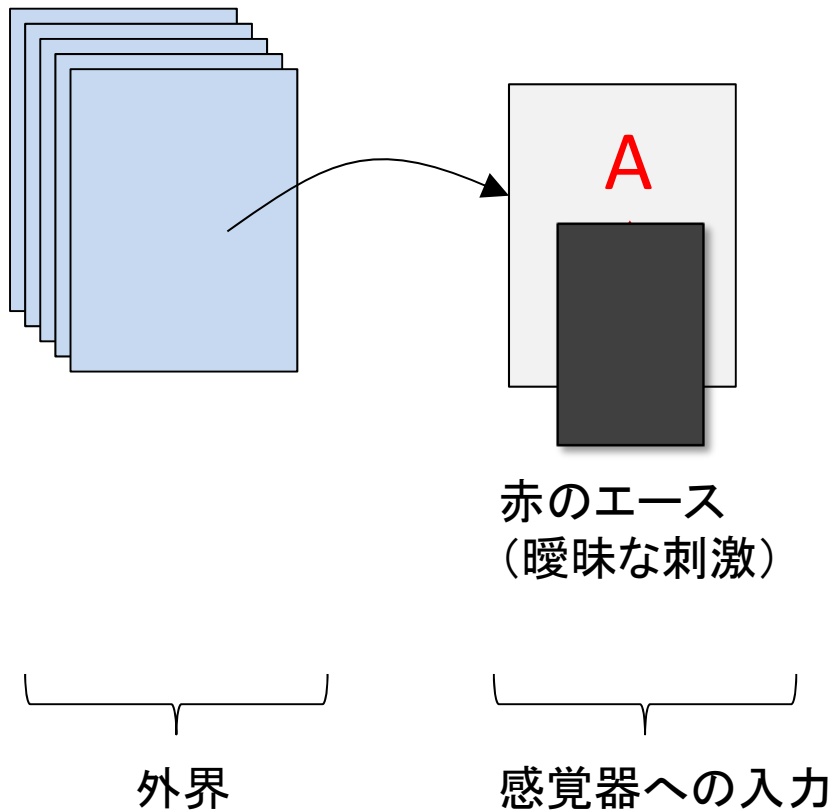


脳

ニューロン活動は事後分布に比例すると仮定



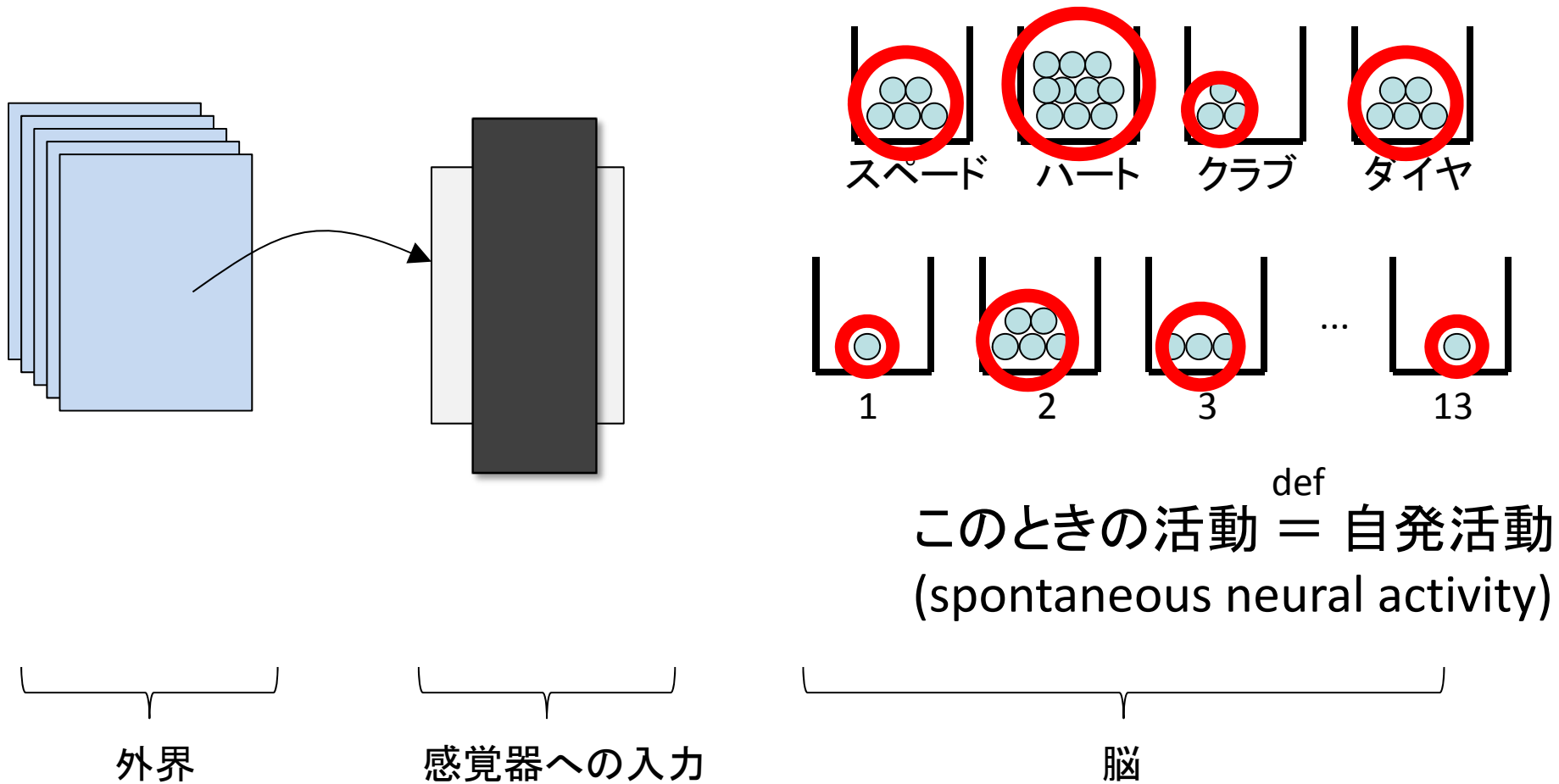
入力が曖昧になると 事後分布はぼやける



過去の統計でハートがダイヤの2倍提示されていれば、ハートがダイヤの2倍の頻度で応答する。

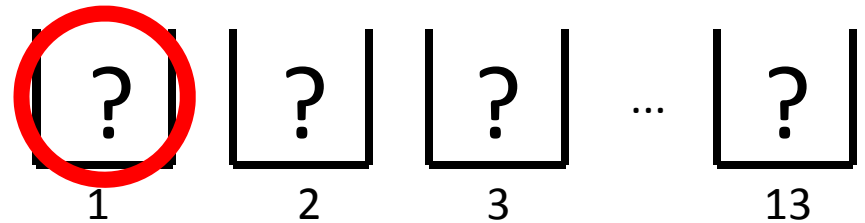
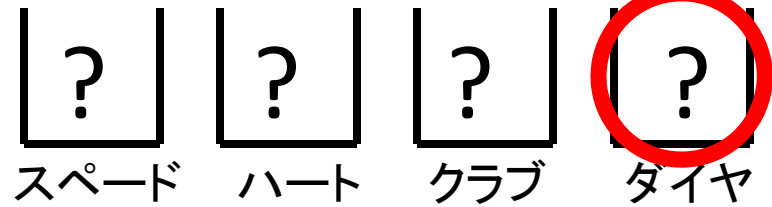
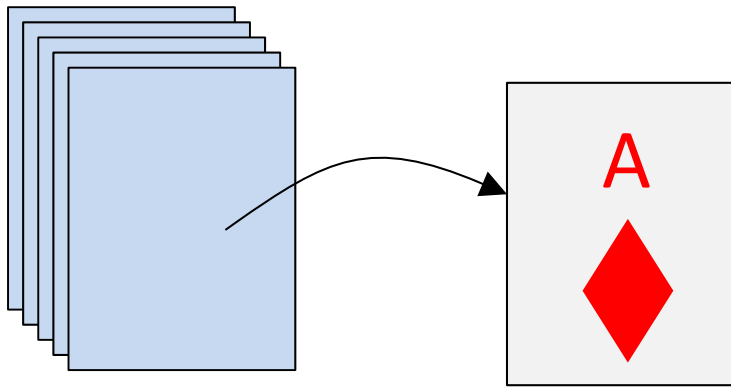
脳

入力がないときのニューロン活動は 事前分布に比例するはず(仮定)

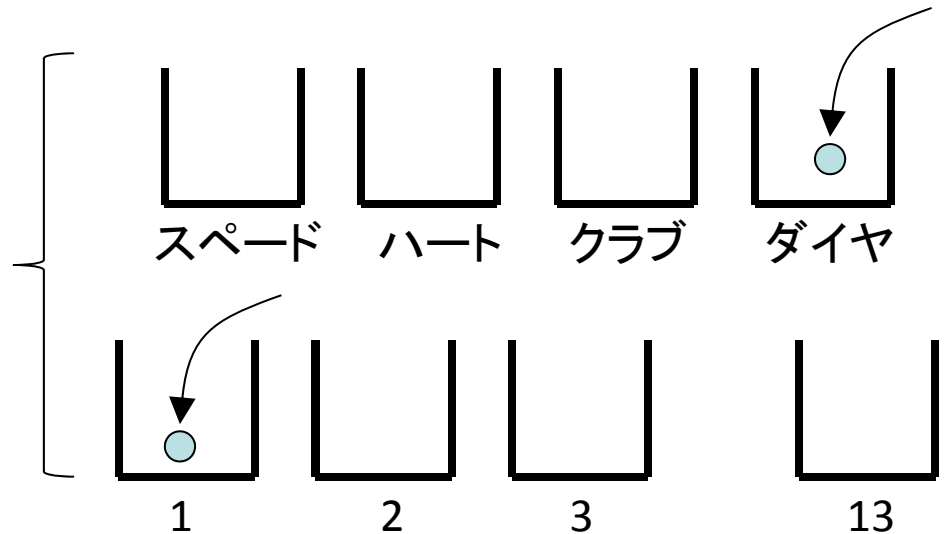


平均誘発活動

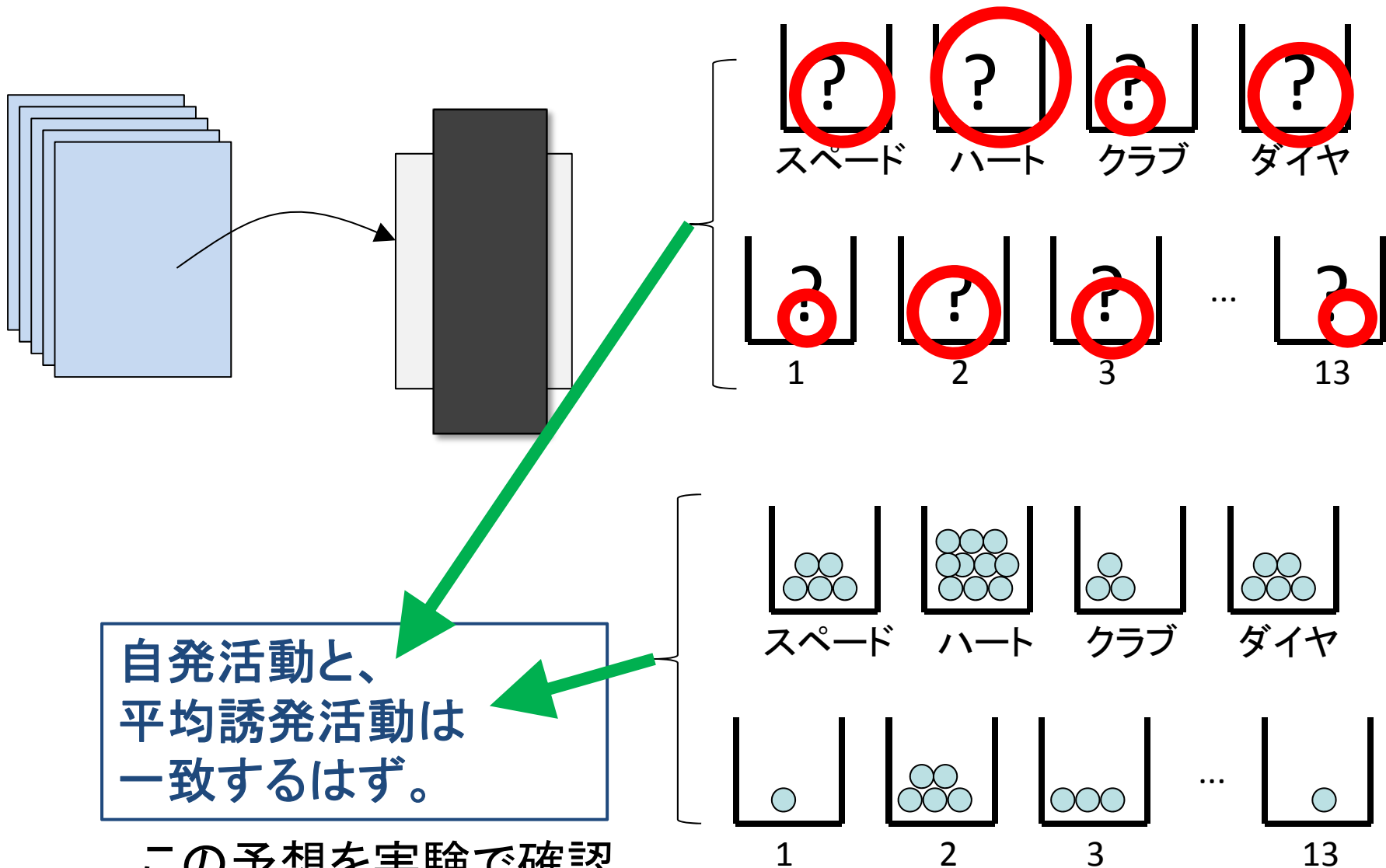
(averaged evoked neural activity)



実際には事前分布そのものは観測不可能。
そこで、自然な入力で誘発されるニューロン活動を計測し、そのヒストグラムを作る(=平均誘発活動)



このモデルからの予想



統計的最適さ(statistical optimality)

についての著者らの主張

Supporting online material P.11-12 の内容を、表現を変えて以下に要約。

h : 特徴 = ニューロン活動

i : 入力

θ : 内部モデルのパラメタ

$P(h|i, \theta) \propto P(i|h, \theta)P(h|\theta)$: 入力から特徴を推定する統計モデル

このような統計モデルが**最適**であるためには下記の条件を満たす必要がある。

(1) 特徴が事後分布で表現される。

(2) 特徴の事後分布が、入力で生成される真の特徴と整合性がある。

このうち(2)を満たすためには例えば

$$\theta_{ML} = \operatorname{argmax}_{\theta} P(i|\theta)$$

(直接確認するのは困難)

や

$$\theta_{ML} = \operatorname{argmin}_{\theta} \operatorname{KL} [P^*(i) || P(i|\theta)]$$

(下記の方法で確認可能)

が成り立てばよい。ただし $P^*(i)$ は真の入力の分布。

仮にKLダイバージェンスがほぼ0、つまり $P^*(i) \doteq P(i|\theta)$ とすると、下記が成り立つ。

$$\int P(h|i, \theta) P^*(i) di$$

(= 自然画像に対する誘発活動の平均)

$$\doteq \int P(h|i, \theta) P(i|\theta) di = \int P(h, i|\theta) di$$

$$= P(h|\theta)$$

(= 活動の事前分布)

よって、**統計的最適なら「平均誘発活動 \doteq 事前分布」が成り立つはず。**

著者らの主張を整理・・・

1. 過去の行動実験データからの推測や一部のモデルの性質から「自発活動＝事前分布」
2. 理論的には、脳が持つ内部モデルが統計的最適ならば「平均誘発活動＝事前分布」
3. もし 1, 2 を真実と認めるならば、
「平均誘発活動＝自発活動」
が「脳が持つ内部モデルが統計的最適になっている」ことを証明する証拠(hallmark)であると言える。この証拠を実験により示す。

補足

- この論文は 1, 2 がほぼ正しいものと仮定して 3 を実験で確認するというストーリーになっている。でも 1 が正しいかどうかは、確立されていない。
- 個人的には 1 よりも 3 の「脳は統計的に最適」の方がよほど確立されているように感じる。なので、むしろこの実験結果は、1 の「自発活動＝事前分布」という仮説を支持する証拠の1つとしても価値がある、と私は受け取った。

以下、論文の内容にそって説明

著者らの予測を説明する模式図

視覚刺激の明るさまたはコントラストを減らせば、視覚皮質における刺激誘発活動は(モデルによっては)事前分布に近づくはず。

Fig.1 A

刺激誘発活動
=事後分布

自発活動
=事後分布=事前分布

様々な年齢の個体で、
自然な動画、
暗闇、
人工的な動画、
に対するニューロン活
動を比較。

Fig.1 B

若い個体：

自然動画での活動 \neq 自発活動

人工動画での活動 \neq 自発活動

大人の個体：

自然動画での活動 \doteq 自発活動

人工動画での活動 \neq 自発活動

となるはず。

実験条件

- 16頭の異なる年齢のフェレットから記録。
 - (育った環境は不明。)
- 200 μ m 間隔の16個の電極、V1 2-3 層。
- 覚醒状態、頭部は固定、目は自由に動かせる状態で視覚刺激を提示。
- 「自然な動画」として見せる movie は "The Matrix" の予告編。

解析手順

Fig.2 A

- ・ 16個の電極から記録。スパイクを 2ms の bin で区切って2値化し、activity pattern を得る。可能な activity pattern の数は $2^{16}=65536$ 。
- ・ 刺激条件ごとにこの activity pattern のヒストグラムをとる。そして、異なる条件の間のヒストグラムのKLダイバージェンスを計算する。

KLダイバージェンス

2つの確率分布の差異を計る尺度。
分布の違いが大きいほど値が大きくなる。

$$D_{\text{KL}}(P\|Q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} dx$$

「カルバック・ライブラー情報量 - Wikipedia」より

成長にともなう変化

M: 自然動画に対する平均誘発活動

S: 自発活動

Fig.2 B

予測通り、成長するにつれて、
平均誘発活動と自発活動の違いが減っていく。

点線は誘発活動のデータの前半と後半の間のダーバージェンス

activity pattern ごとの頻度でプロット

Fig.2 C

成長後は、各パターンの発生頻度が平均誘発活動と自発活動で等しくなっている。

成長後は平均発火率は上がっている。
(スパース性が減っている。)

M: 自然動画に対する平均誘発活動
S: 自発活動

分布のどの側面が影響して2つの分布が近づくのか

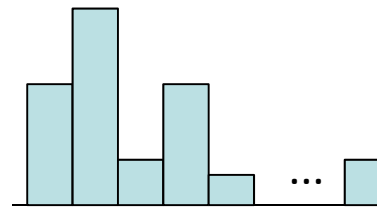
- 多くのV1の計算論的モデルが仮定するように、学習によって「情報表現の冗長性が削除される」と仮定すると、学習が進むにつれスパース性は増し、情報チャンネル間の独立性も増すと予想される。その影響ではないのか？
- この点を確かめるために「サロゲート分布」(ある側面だけ取り除いた人工データ)と比較

サロゲート分布

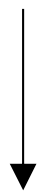
計測データ(各チャンネルは独立とは限らない)

010001101
100110110
000100001
...
001011000

→
ヒストグラム作成



事後分布

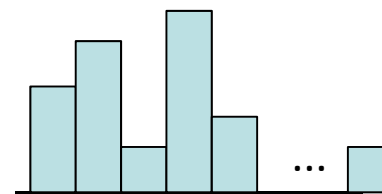


各チャンネルが独立だと仮定し、
各チャンネルの発火頻度が同じ人
工データ(空間的相関を取り除い
たデータ)を生成

サロゲートデータ

000100011
100010001
001101010
...
110101000

→
ヒストグラム作成



サロゲート分布

もし各チャンネルが
独立なら2つの分布
は一致

サロゲート分布との比較

Fig.3 A, B

M, \tilde{M} : 自然動画に対する平均誘発活動とそのサロゲート分布

S, \tilde{S} : 自発活動とそのサロゲート分布

成長とともにチャンネル間の
独立性は減る。

自発活動の空間的相関を
取り除いた分布は平均誘
発活動に近づかない。

推移確率分布と時間的相関を取り除いたサロゲート分布との比較

Fig.3 C,D

τ ms経過による応答パターンの推移確率分布の間のKLダイバージェンス
(adult P129 to P151)

20ms間隔以内の活動には相関がある。

自発活動の時間的相関を取り除いた分布も平均誘発活動に近づかない。

人工動画における活動との比較

Fig.4 A,B,C

M: 自然動画
N: 白黒正方形のノイズ
G: 動く縞模様
S: 自発活動

活動分布をMを原点にして、
多次元尺度構成法で表現。
色の濃さは年齢を表す。

N, G はMほどはSに近づかない

Sは成長とともにMに近づくが、N, Gは近づかない。

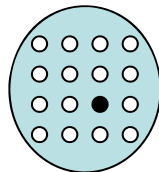
M, N, G どちらに成長とともに近づかない。

実験結果のまとめ

- 成長とともに平均誘発活動は自発活動に近づく。
- 成長とともに、スパース性が減り、チャンネル間の独立性は減る。
- 20ms以内の発火パターンの時間的相関も、成長とともに平均誘発活動と自発活動とが近づく。
- 人工動画は自然動画ほどは、自発活動に近づかない。

個人的感想1

- 私は今まで事前分布は一様分布になってノイズと区別がつかないだろうと思っていた。この論文の結果はそうではなく、意外。
 - 電極間隔 $200\mu\text{m}$ はマクロコラムより小。競合学習により排他的に活動するようになったのかも。



個人的感想2

- 「成長とともにスパース性が下がる」ことを知ったのは収穫。
 - (それを真似すれば学習アルゴリズムの汎化性能を上げられるかも。)

個人的感想3

- 内部モデルに対して制約がほとんどない。
 - 著者はそれが利点だと言っているが、脳の内部モデルの具体的な形が知りたい私にとっては新たなヒントが得られなかったので残念。

個人的感想4(重要)

- この論文では「事前分布」と「注意」との関係に全く触れていない。
 - 「視覚刺激提示前に注意の影響を受けるニューロン」はV4に少しあるもののV1, V2では見つからない(*)。しかし、この論文同様の暗闇実験によってV1から見つけることができるかもしれない。もし見つければ、「事前分布=注意」という仮説の強い証拠となるはず。

(*) A backward progression of attentional effects in the ventral stream (Buffalo EA, Fries P, Landman R, Liang H, Desimone R, Proc Natl Acad Sci U S A. 2010 Jan 5;107(1):361-5)

<http://www.ncbi.nlm.nih.gov/pubmed/20007766>

Q & A 1

Q:「自発活動＝事前分布」という解釈以外の可能性はないのか。

A:あり得ると思う。

が、例えば [Weiss et al. 2002] の実験結果は、コントラストが小さくなるにつれ振る舞いが事前分布に近づくことを示しているので、そこからの類推で「自発活動＝事前分布」は妥当かも。

Q & A 2

Q: 暗闇にしなくても fixation point 以外の場所に受容野があるニューロンの自発活動に注意の影響(事前分布)が出るのでは。

A: V1, V2では見つからなかったと言う報告がある(*)。目の前に fixation point しかないと高次の領野が認識したら、その文脈情報が事前分布としてV1に降りて、fixation point 以外のすべての特徴の事後分布が0になるのだと思う。

以上

コメント、ご質問、間違いの指摘などを歓迎いたします。

y-ichisugi@aist.go.jp