

視覚を理解し応用する  
～工学・理論・生理・心理のいいとこ取り～

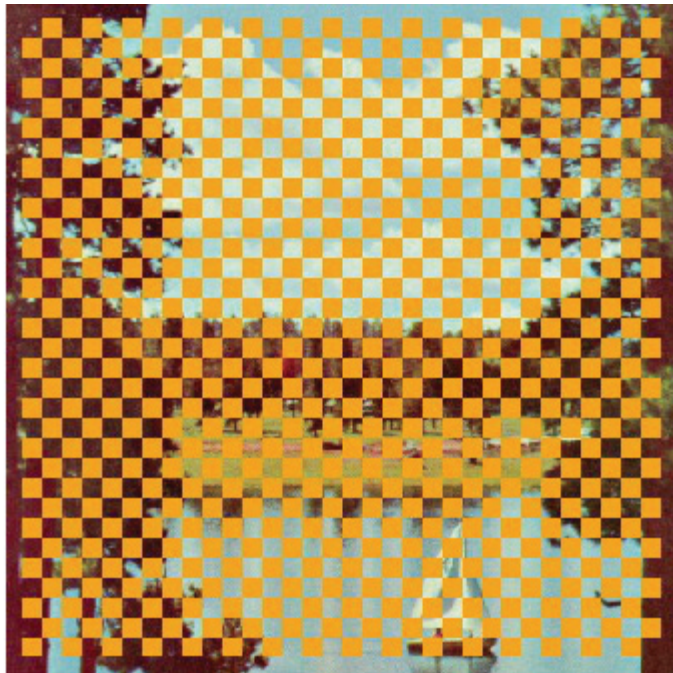
情報システム学研究科  
佐藤俊治

# 勝手に自己紹介

- 名前: 佐藤俊治 (さとうしゅんじ)
  - よくある名字、佐藤 と覚えてください。
- 経歴など
  - 1973年: 岩手県に生まれたのち、札幌、群馬、再度岩手
  - 2000年: 東北大学大学院工学研究科修了 博士(工学)
  - 2000年~2001年: 日本学術振興会PD
  - 2001年~2004年: 東北大学大学院工学研究科助手
  - 2004年~2006年: 東北福祉大学常勤講師
  - 2006年~2009年: 理化学研究所BSI研究員
  - 2009年~現在: 国立大学法人・電気通信大学大学院准教授
- 受賞歴
  - 2006年: 日本神経回路学会奨励賞
  - 2009年: 日本神経回路学会論文賞
  - 2010年: APNNA Young Researcher Award

## 本日の講演内容 2つのうち1つ

- 神経生理学・計算論・  
画像工学の融合例



- 認知心理学・計算論・  
画像工学の融合例



# 研究そのものの背景

# 画像処理研究をしておりました

- 文字認識
- 物体認識
- 画像理解
- 画像生成



SEIUN : 高精度高速知的文字認識システム

大ショック ( ; °° )

# 視覚研究を始めた理由

1. 「認識率が良い・悪い」を評価しているのは何か？
2. ヒトの視覚系(脳)である
3. 脳の仕組み・動作原理を参考にすれば、究極の画像処理システムができるはず
4. こ、こ、これだ！  
(短絡思考)



# 工学部出身佐藤が考える 視覚計算論モデルが満たすべき要件

## 1. 科学的要件(当たり前):

- 現象・データを数理的に解釈し、
- 実験結果を再現し、
- 複雑な神経特性をすっきりと説明する数理モデル

} 当たり前

## 2. 工学的要件(さらに加えて):

- 数理モデルはそれ自体が有用な画像処理アルゴリズムであり(あってほしい)
- 既存の画像処理手法との対応・比較ができる数理モデル

} 最近ほとんど見かけない

## 3. なぜならば、

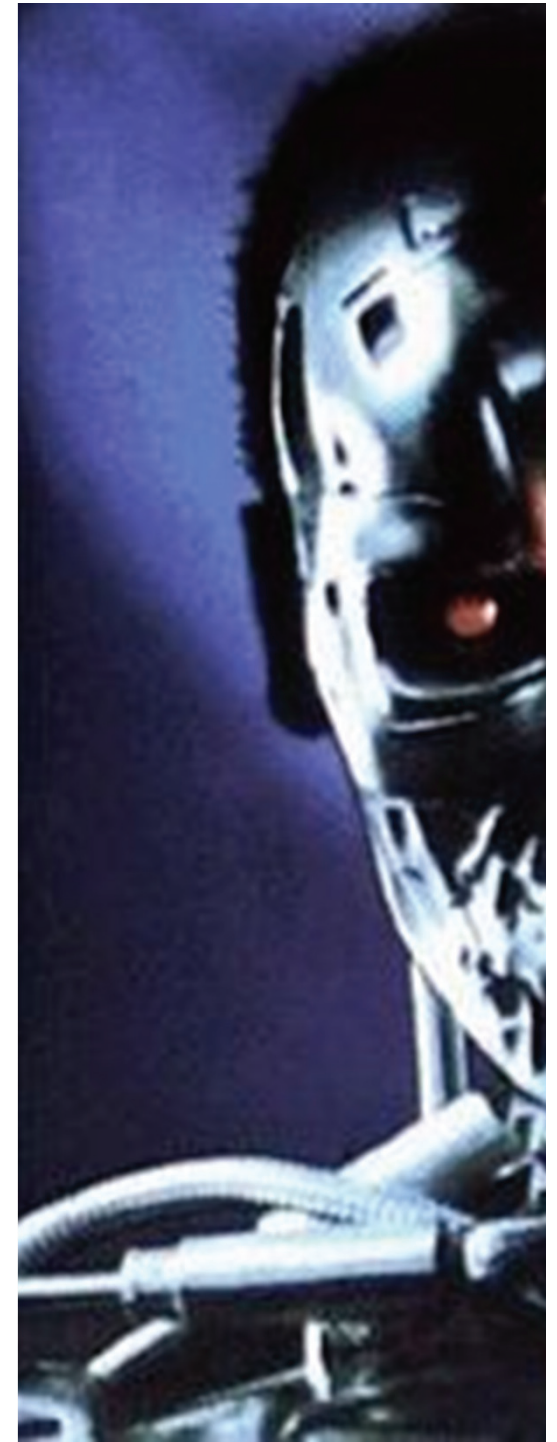
- 視覚脳科学の対象は「中身や動作原理はサッパリよくわからないけれども、超高性能画像処理マシン=視覚」であり、すなわち
- 視覚脳研究においては、科学的対象と工学的対象が同じであるから

## 4. (2.に戻る)



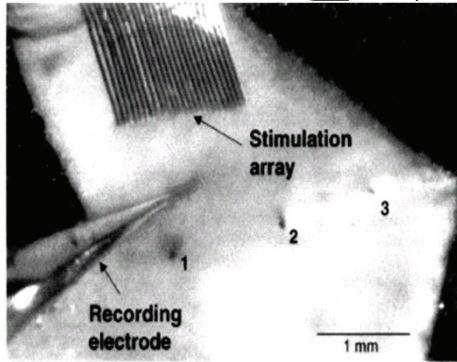
# 視覚 || 超高性能 画像処理マシン

と割り切って考えれば  
視覚脳科学と画像工学の  
対象と興味はほぼ同じ

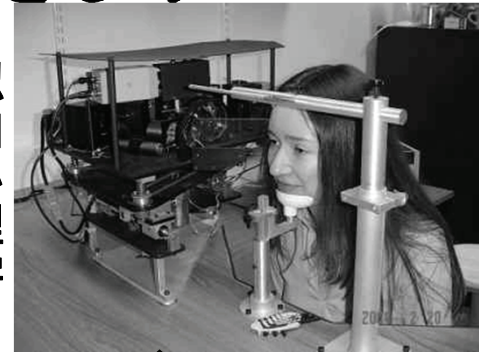




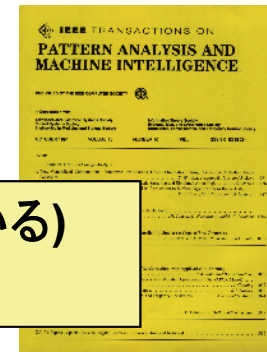
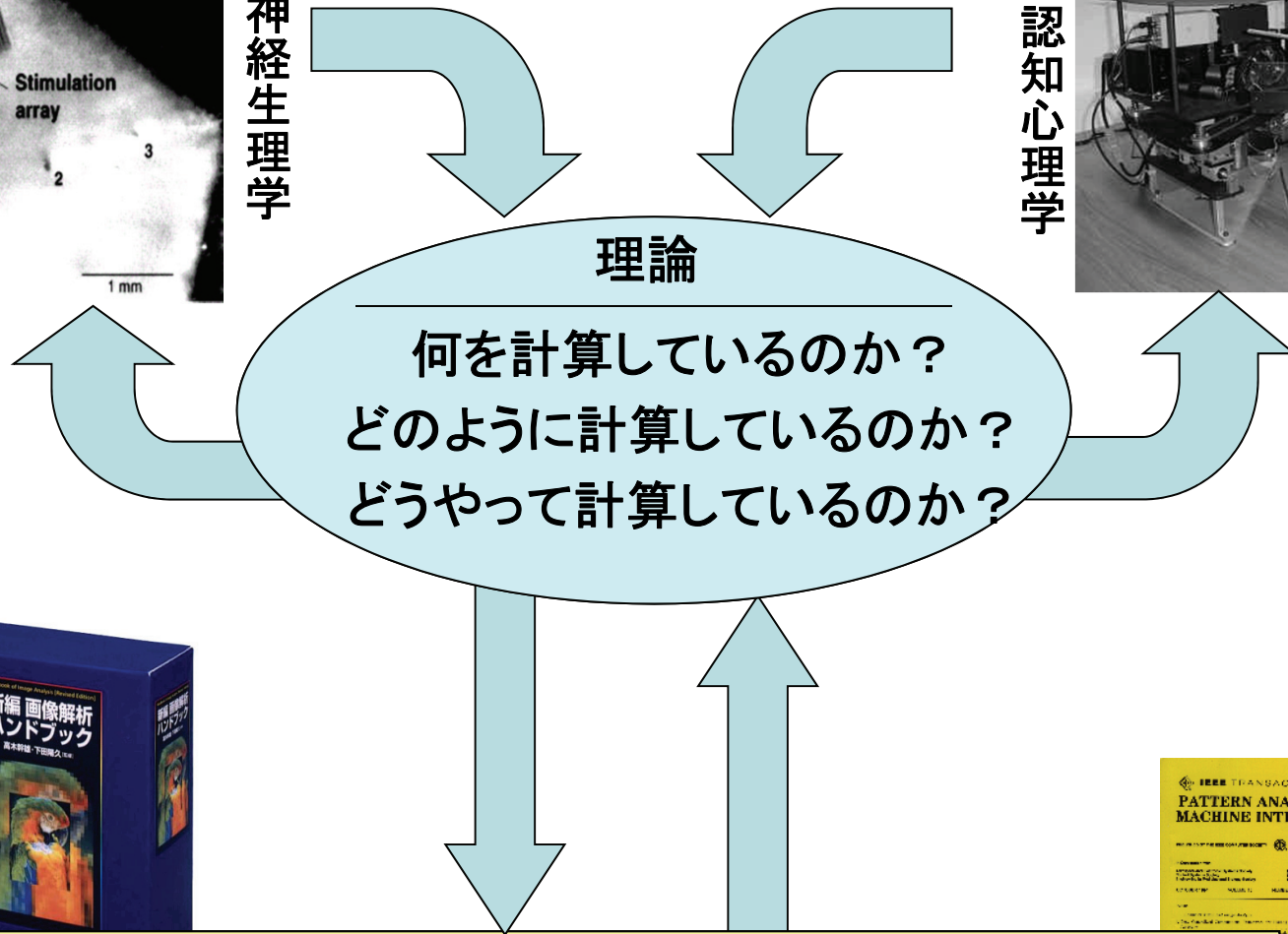
# 特に視覚は、理学でもあり工学でもあるので どの方面からもアプローチできます



神経生理学



認知心理学

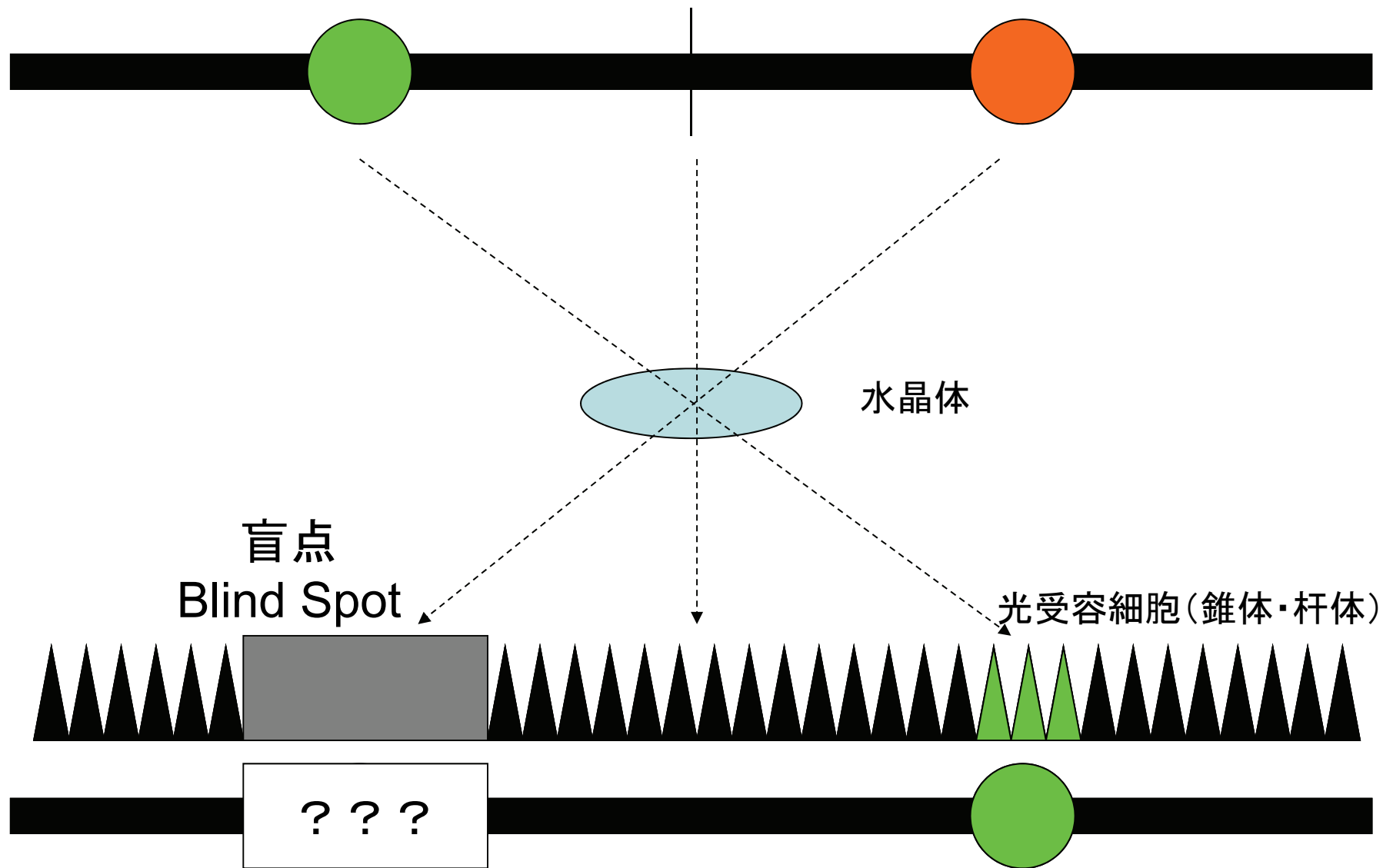


(視覚特性や細胞特性についてはとりあえず無視している)  
「目標指向的」工学的画像処理

本日のターゲット①:  
盲点における補完(filling-in)

- 生理実験データの意味を解釈し
- 数理モデルを構築し
- 画像工学アルゴリズムとして応用

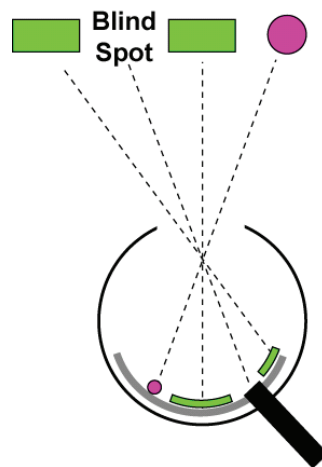
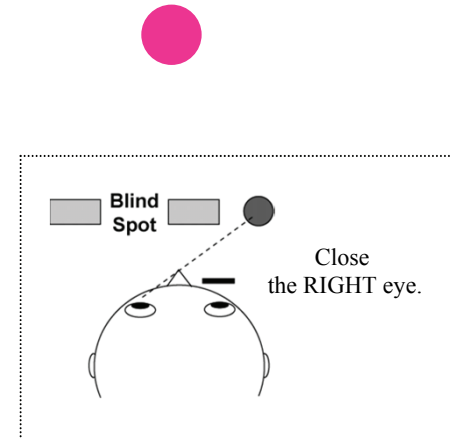
# 盲点における充填知覚



# 盲点における充填知覚を体験してみる

## 盲点

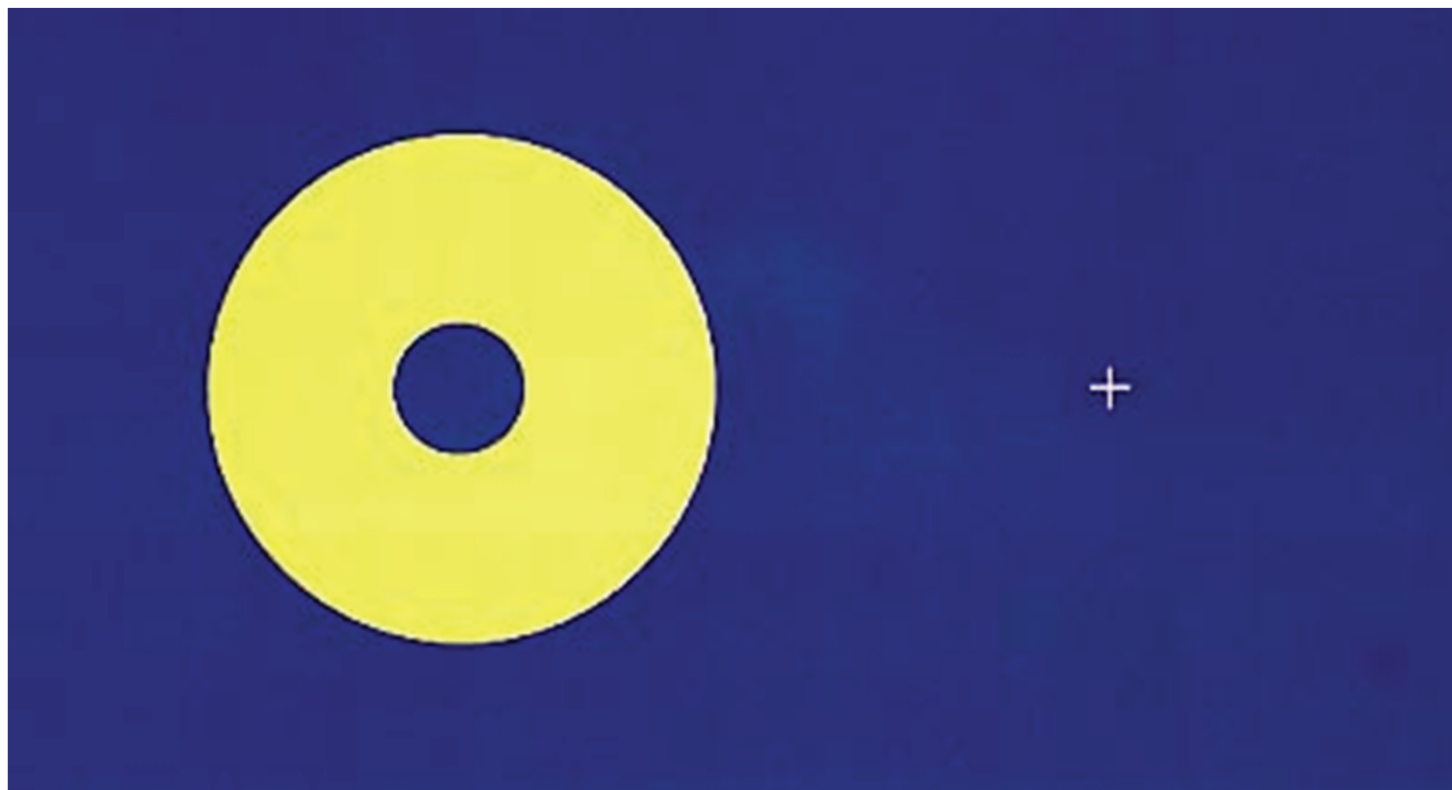
1. この紙を両手で持ち、腕を伸ばしてください。
2. 右目を閉じてください。
3. 紙は左右にずらさずまっすぐ持って、
4. 左目でピンクの○を見てください
5. そのままゆっくり、紙を手前に近づけてください。
6. ある距離まで近づけると、「盲点」の文字が消え、下図のような、途切れていない緑の横棒が知覚されます。



The sheet of photoreceptors is much like a sheet of film at the back of a camera. But it has a hole in it. At one location, called the optic nerve head, processes of neurons collect together and pass as a bundle through the photoreceptor sheet to form the optic nerve (the thick black line extending down and to the right in the diagram), which carries information from the eye to the rest of the brain. At this location, there are no photoreceptors, and hence the brain gets no information from the eye about this particular part of the picture of the world. Because of this, you should have a "blind spot" (actually two, one for each eye), a place pretty much in the middle of what you can see where you can't see.

(adapted from <http://serendip.brynmawr.edu/bb/blindspot1.html> )

## 盲点における充填知覚



<http://www.nips.ac.jp/guide/2002/res/bio-system.html>

右目をとじ、左目で十字をみながら顔を絵から20cmくらいに近づけると、青い内円が消えると同時に円全体が黄色い満月のように見える(充填知覚)。

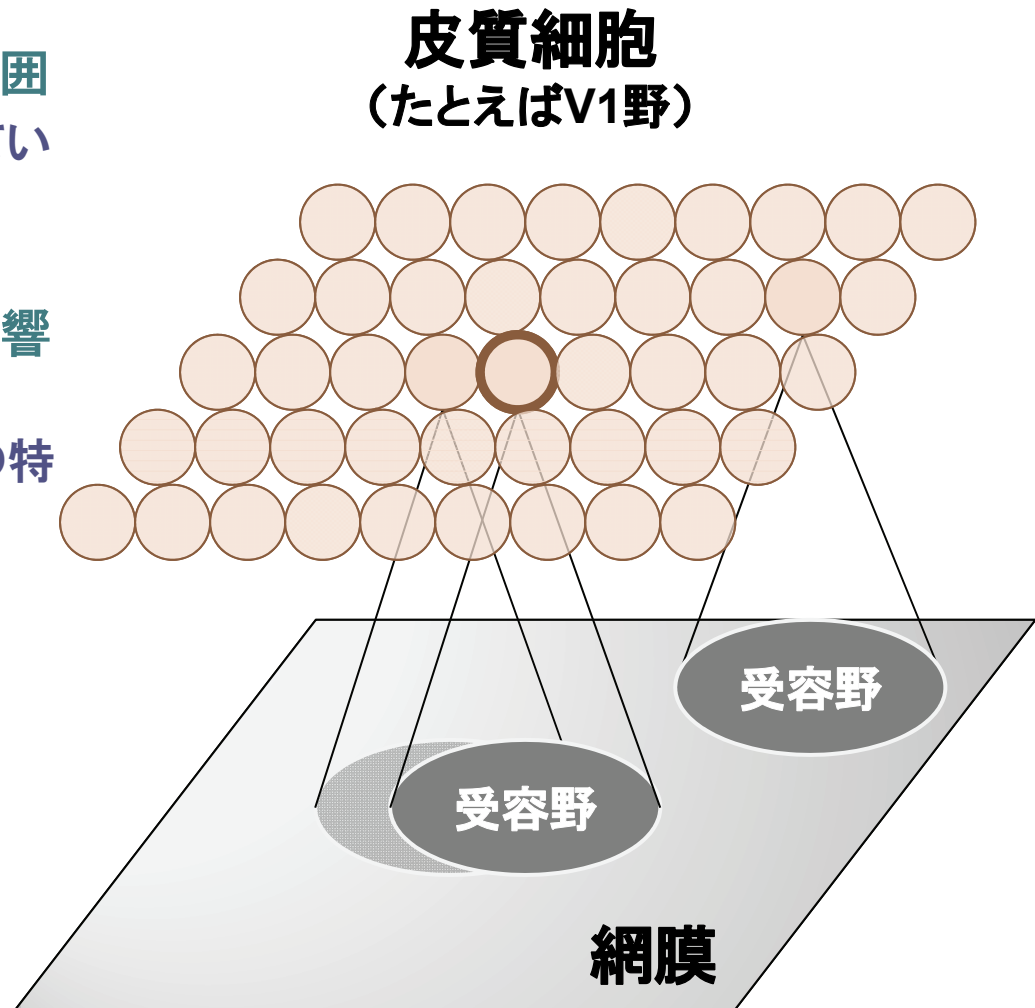
脳内における盲点補完の仕組みが分かれば、  
Image Inpainting (画像修復)のためのアルゴリズムが  
作れるのではないか？



盲点における神経生理学的特性を考慮した数理モデル  
||  
Image Inpainting アルゴリズム

# 受容野

- 受容野
  - 視覚関連細胞1つ1つの処理範囲
    - ・ 網膜像(視野全体)を処理していない
    - ・ 網膜像の一部だけを処理
  - 「受容野」は細胞の活動度に影響する「網膜の局所範囲」
    - ・ 受容野の空間構造が、細胞の特性を決める



*J Neurophysiol* 93: 2374–2387, 2005.

First published January 5, 2005; doi:10.1152/jn.00811.2004.

## Neural Responses in the Macaque V1 to Bar Stimuli With Various Lengths Presented on the Blind Spot

**Masayuki Matsumoto and Hidehiko Komatsu**

*Division of Sensory and Cognitive Information, National Institute for Physiological Sciences; and Department of Physiological Sciences, The Graduate University for Advanced Studies, Okazaki, Aichi, Japan*

Submitted 9 August 2004; accepted in final form 31 December 2004

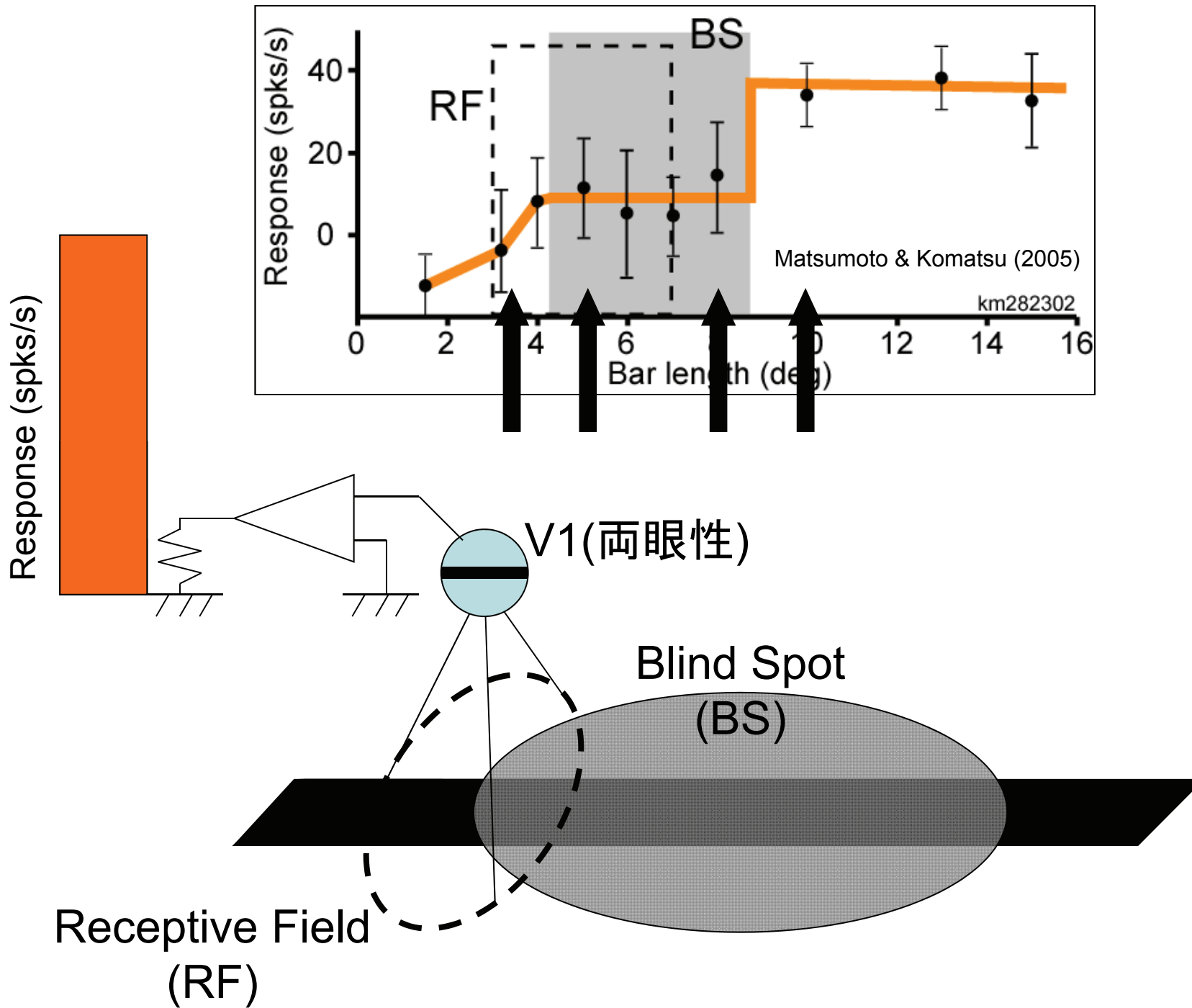
# 本研究の基礎となる 神経生理学的知見の紹介 (具材の性質)

Matsumoto & Komatsu (2005)

*J. Neurophysiology*

**93:2374--2387**

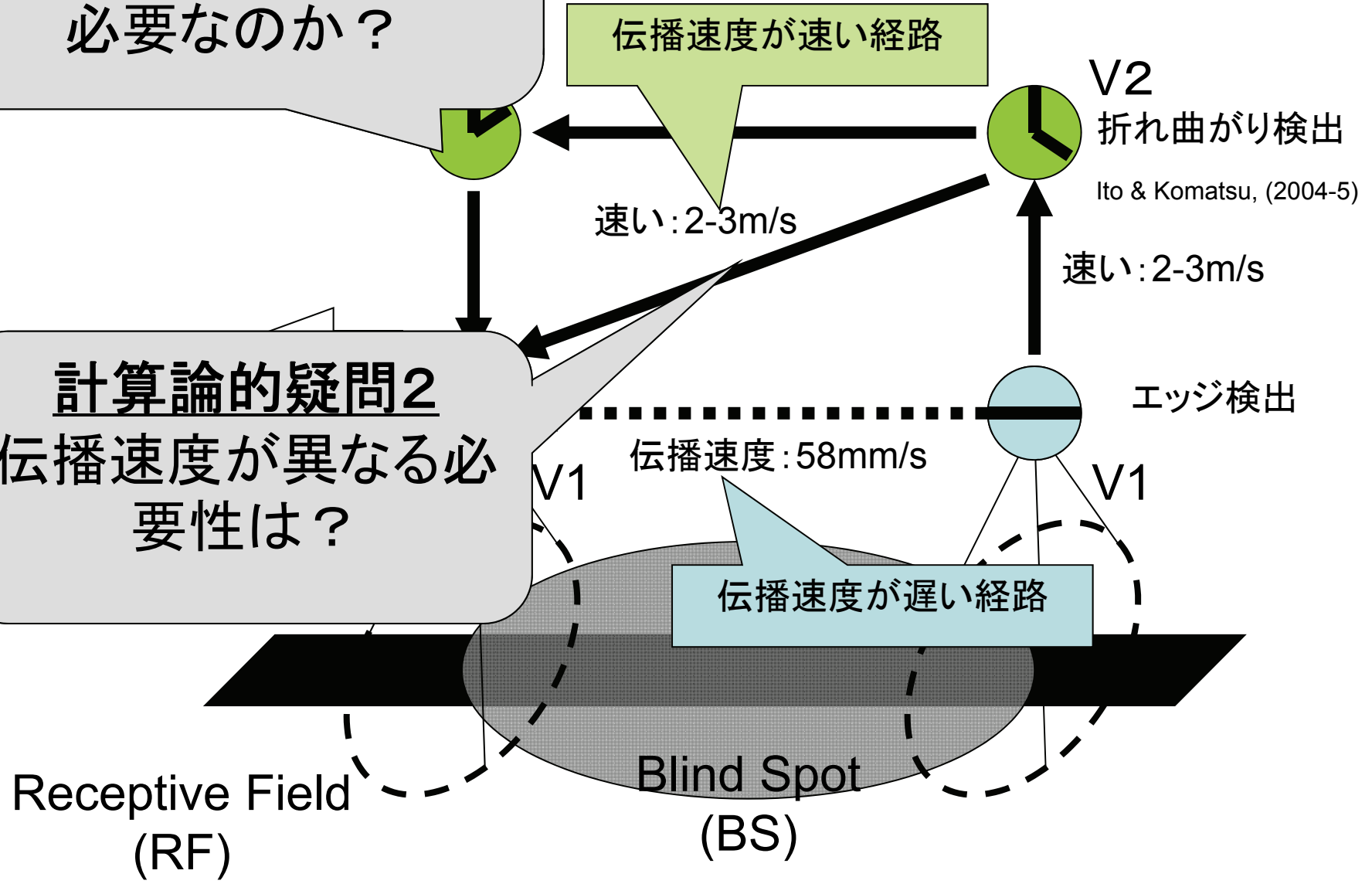




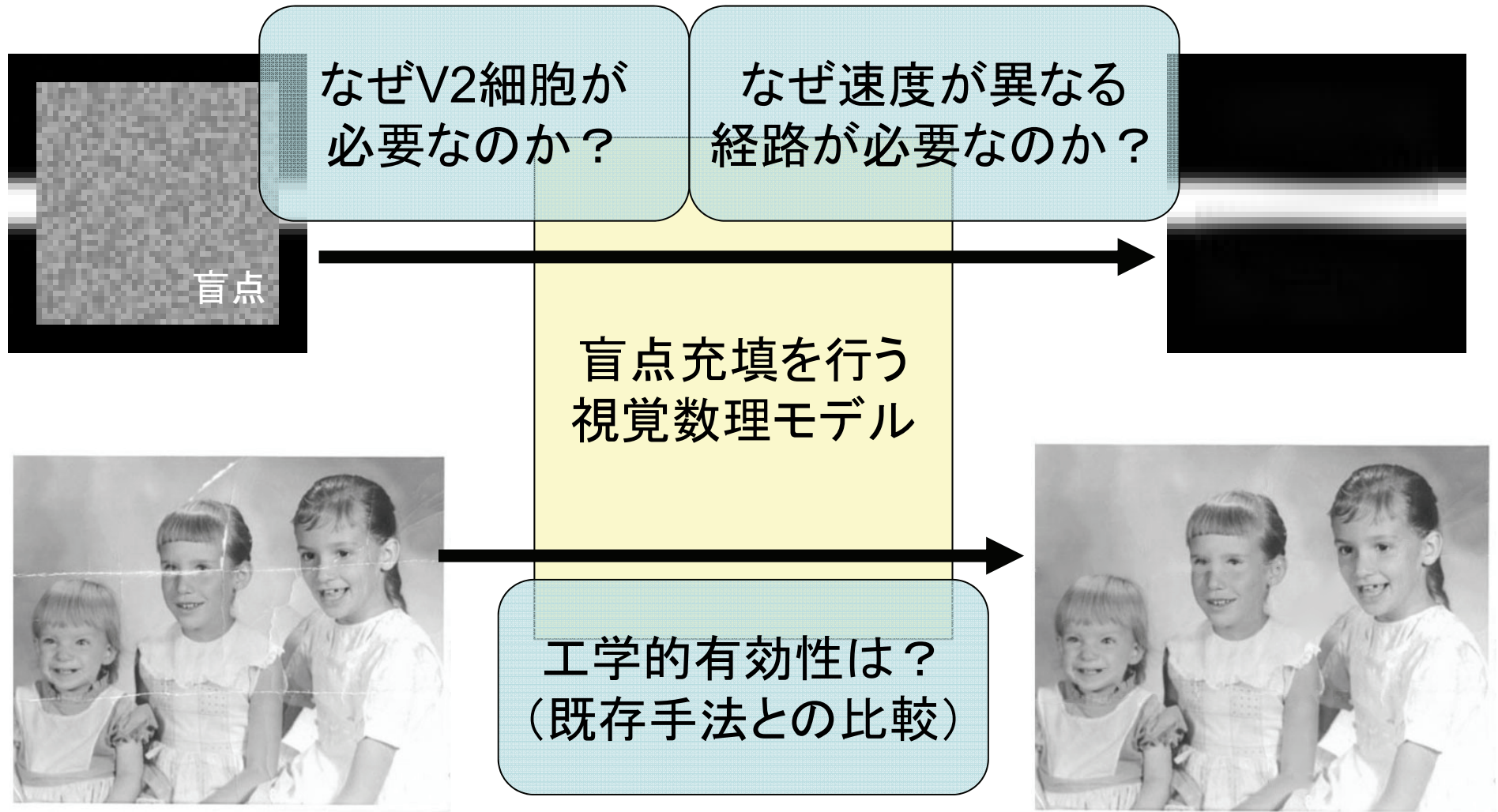
# Komatsu の概念的モデル

**計算論的疑問1**  
なぜV2細胞の情報が  
V1細胞での充填に  
必要なのか？

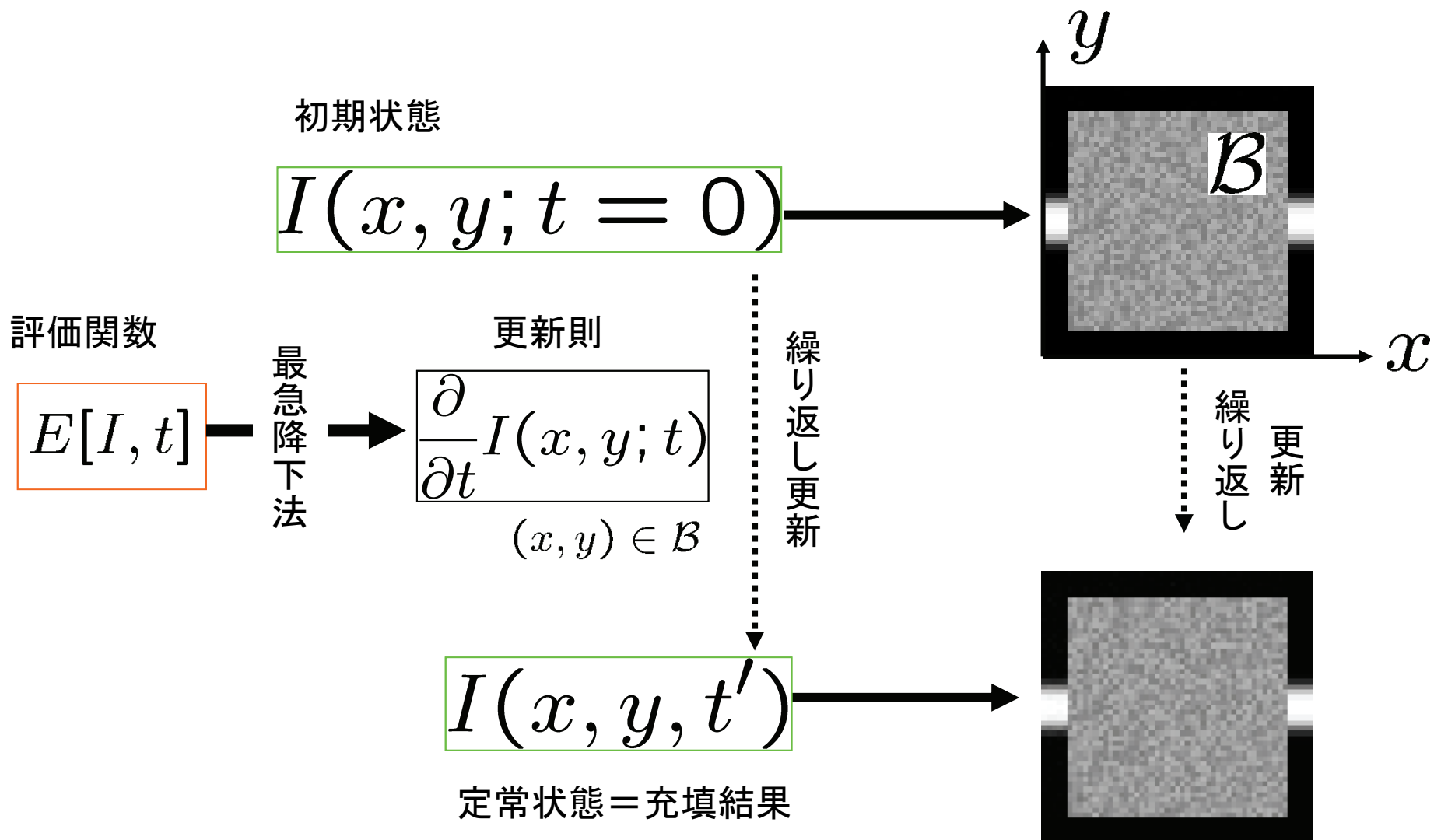
**計算論的疑問2**  
伝播速度が異なる必  
要性は？



# 本研究の目的



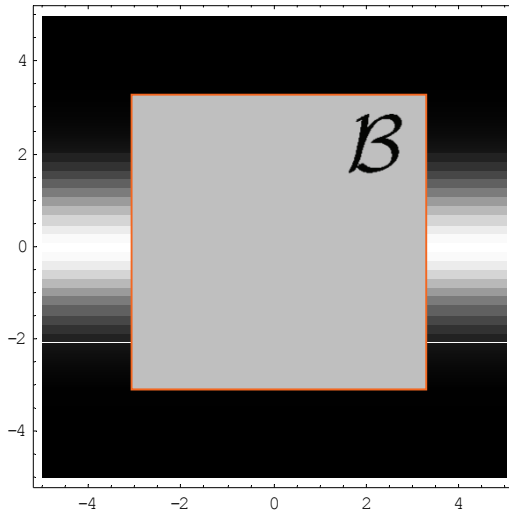
# 充填のためのアルゴリズム



# 評価関数 (a.k.a. エネルギー, 汎関数)

$$E[I, t]$$

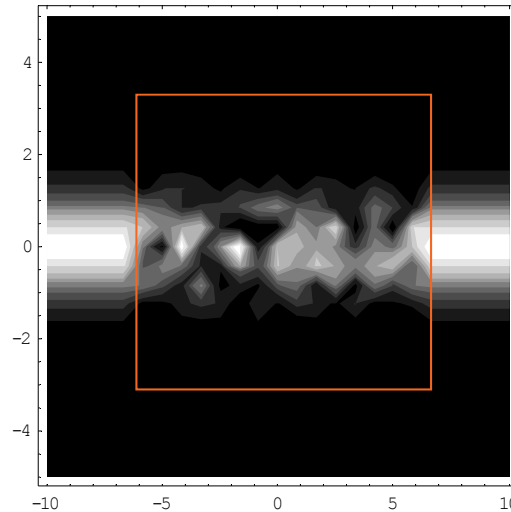
時刻  $t$  における充填画像  $I$  が、望みの画像なのか、  
そうでないのかを定量評価する関数



初期状態

$$E[I, t = 0]$$

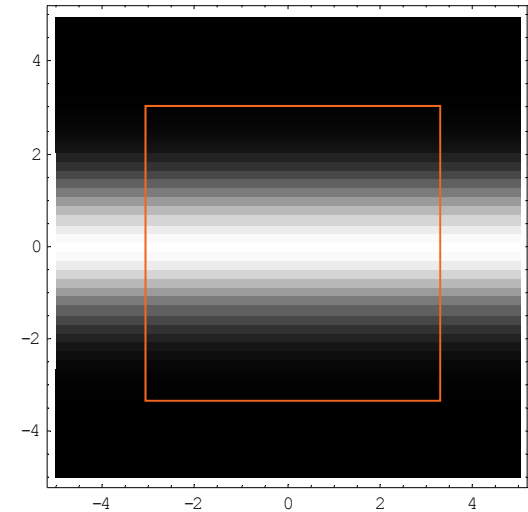
=大



望ましくない画像

$$E[I, t_1]$$

=大



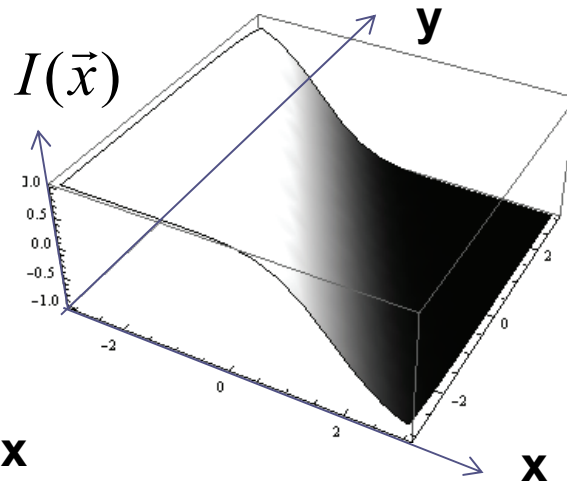
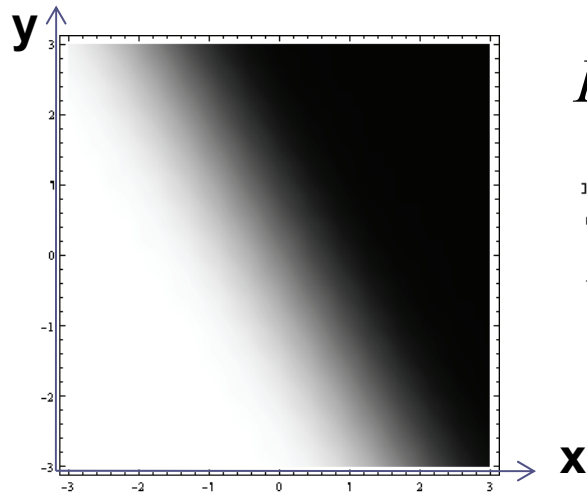
望ましい画像

$$E[I, t_2]$$

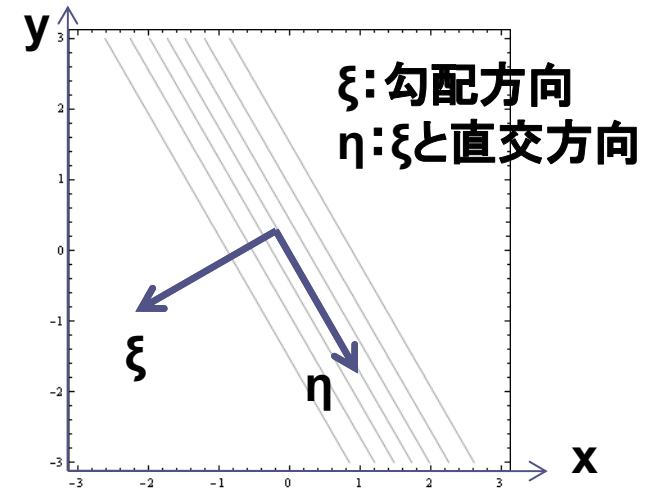
=小

# 画像の表現・局所座標系 $(\xi, \eta)$

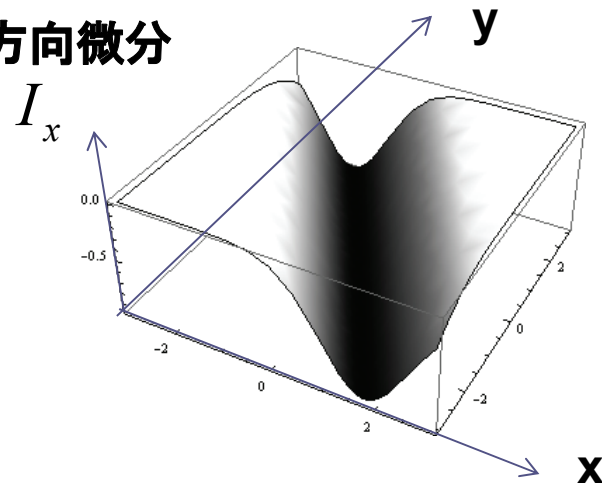
網膜像  $J(\vec{x}), I(\vec{x})$



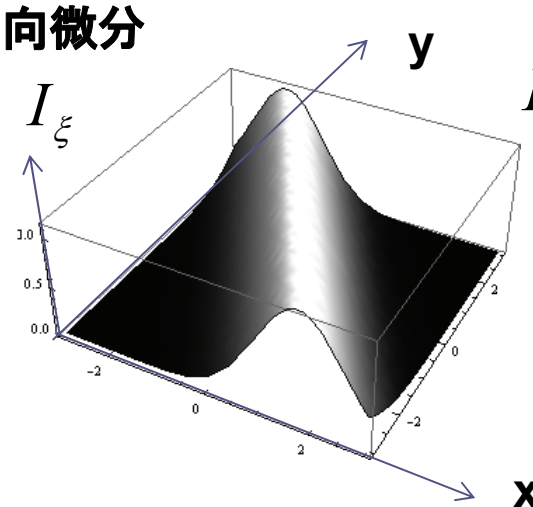
値の等高線=レベルセット



方向微分



方向微分



$$I_{\xi} = \|\nabla I\|$$

$$= \sqrt{I_x^2 + I_y^2}$$

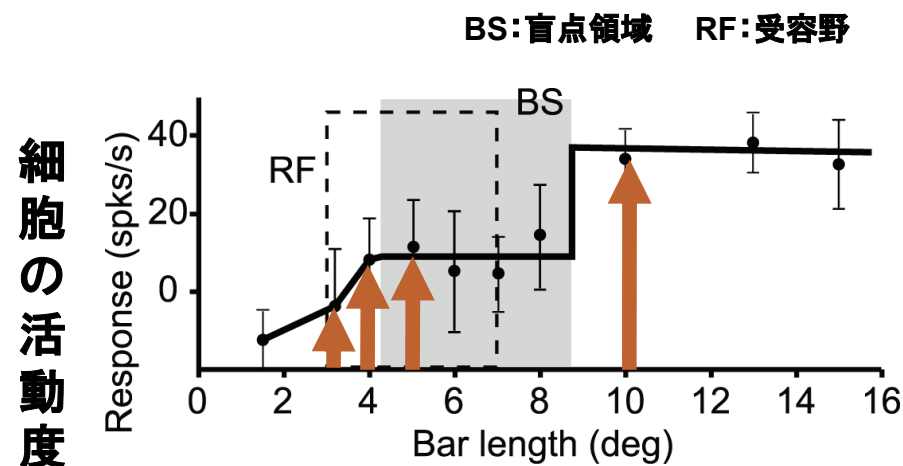
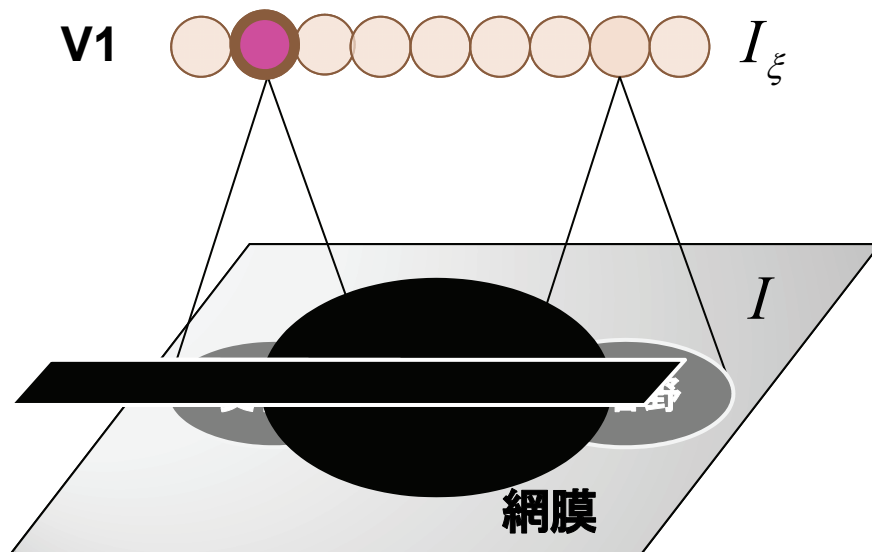
=エッジの強度

= (V1細胞でコーディング)

# 盲点補完はどこで行われているのか？

Matsumoto & Komatsu (2005) J. Neurophysiology

- V1細胞で既に補完(充填; filling-in)が行われている



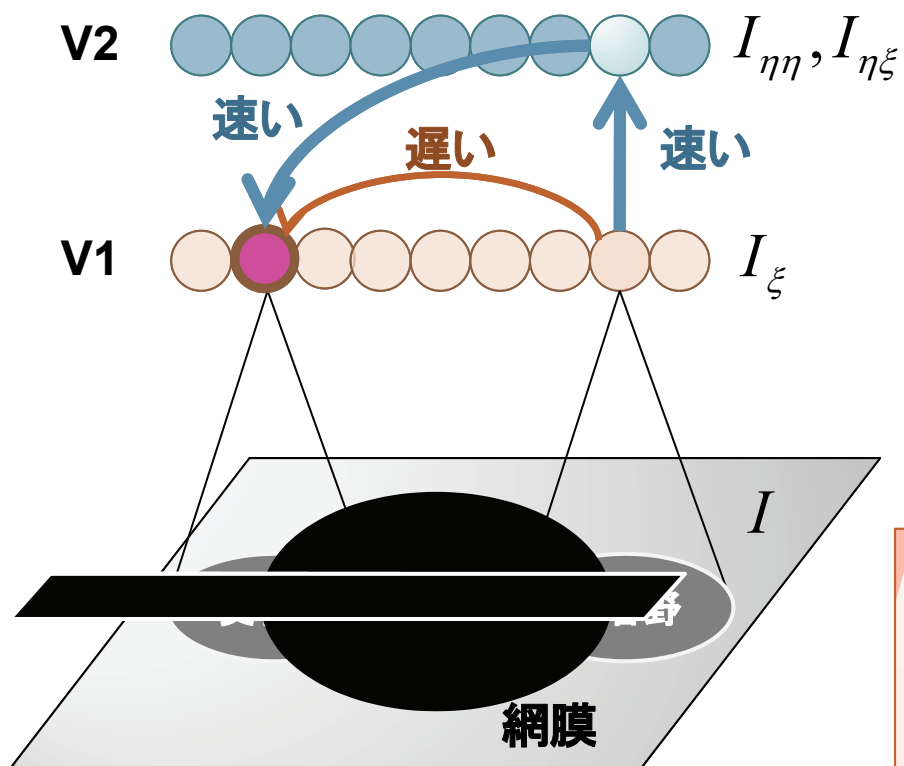
提示刺激であるバーの長さ

あたかも、盲点領域など最初から存在せず、一本の長いバーに対するような反応を示す

# 盲点補完はどこで行われているのか？

Matsumoto & Komatsu (2005) J. Neurophysiology

- V1細胞で既に補完(充填; filling-in)が行われている



1. V2細胞の介在がないと説明できない現象あり
  - V2細胞はより複雑な画像特徴に対して反応
  - 例: 曲率 (Ito & Komatsu (2002))
2. 2つの経路の信号伝播速度が異ならないと説明できない現象あり
  - 遅い: V1  $\rightarrow$  V1
  - 速い: V1  $\rightarrow$  V2  $\rightarrow$  V1

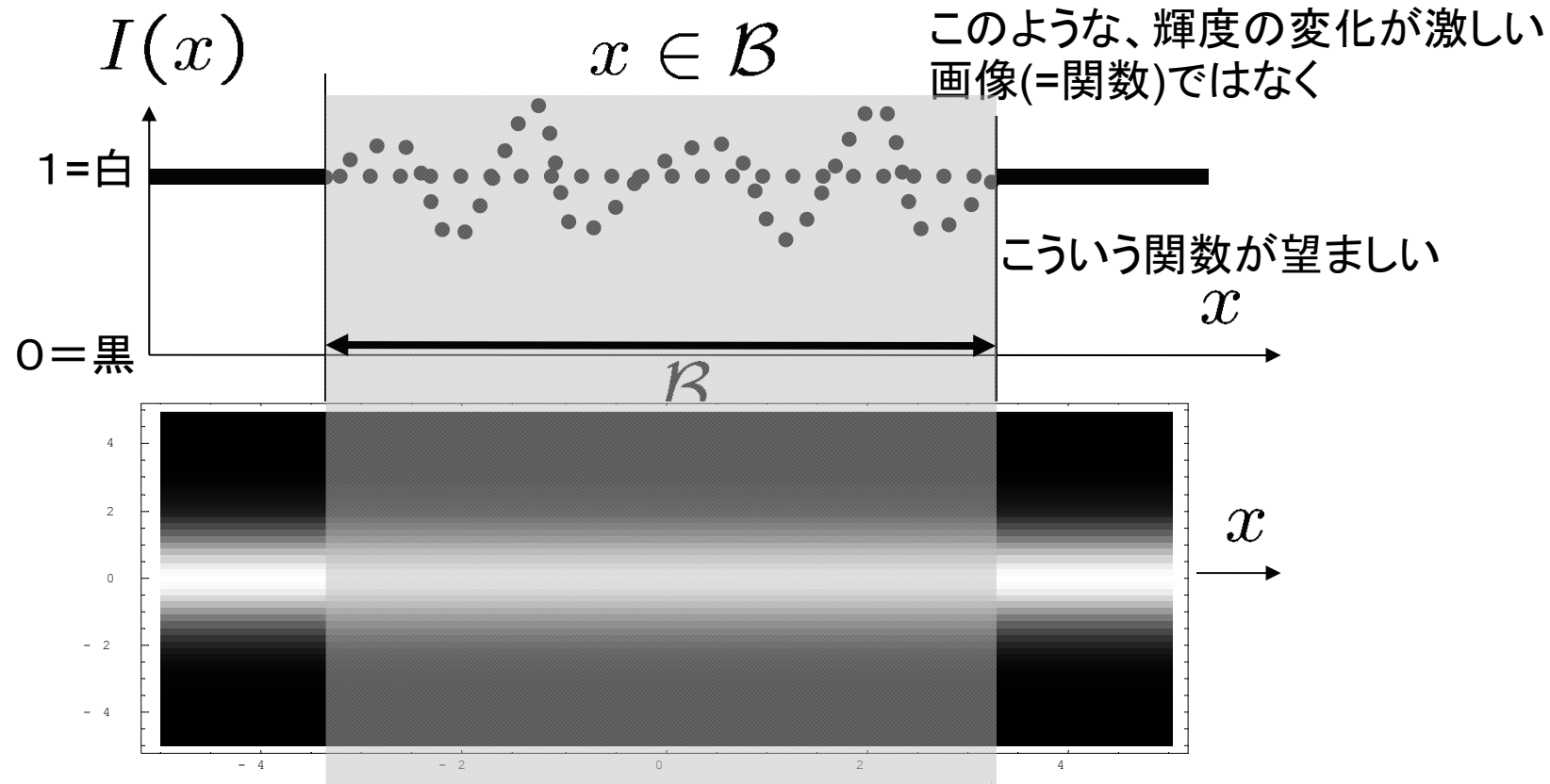
**目的** これら神経特性の意味を理論的に明らかにし、モデルを構築する。

1. V1における充填に、なぜV2が必要？
2. なぜ速度が異なる必要がある？



# 評価関数を考える

(望ましい充填画像を式で表現する)



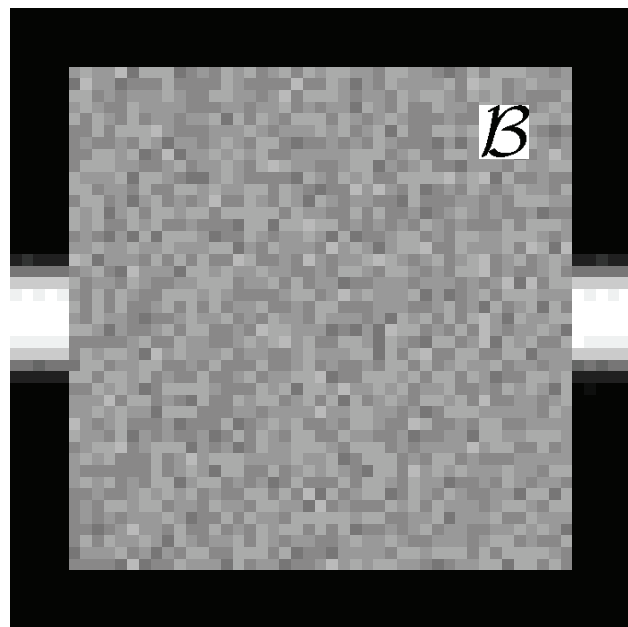
$$\forall \vec{x} \in B \quad \left\| \frac{\int_B \nabla I(x) dx}{|B|} \right\| \rightarrow 0$$

V1細胞がコードしている情報 : エッジ方位・強度検出結果

# V2細胞がコードしている情報も考慮すべし

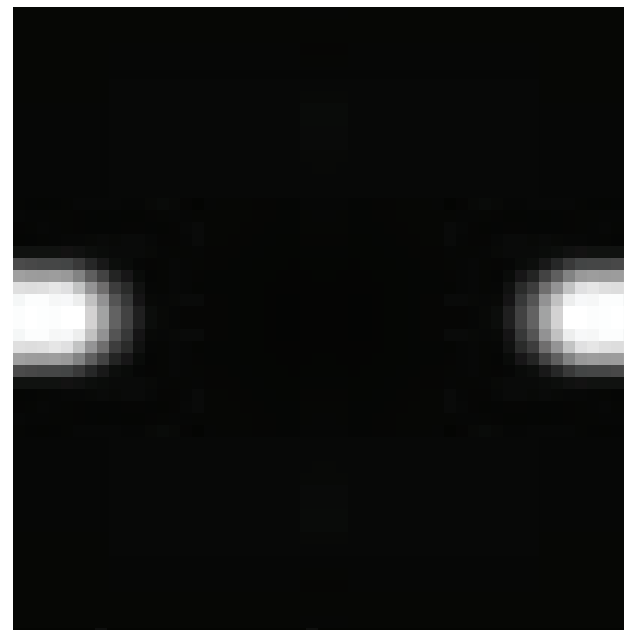
$$E[I] = \int_{\vec{x} \in B} I_{\xi}^2(\vec{x}, t) d^2\vec{x}$$

V1細胞がコードしている情報 : エッジ方位・強度検出結果



$I(\vec{x}, t = 0)$

$$\frac{\partial}{\partial t} I(x, y; t)$$



$I^*(\vec{x})$  理想

# V2細胞がコードしている情報も考慮する

Behavioral/Systems/Cognitive

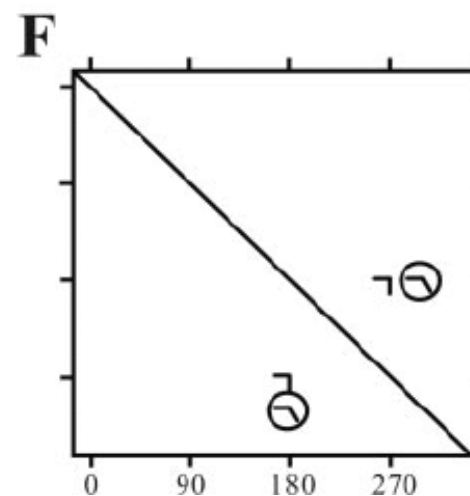
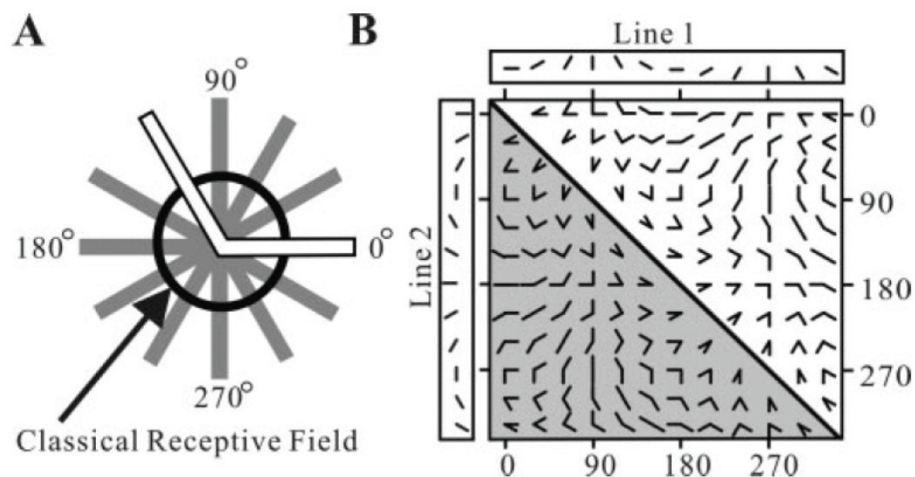
The Journal of Neuroscience, March 31, 2004 • 24(13):3313–3324 • 3313

## Representation of Angles Embedded within Contour Stimuli in Area V2 of Macaque Monkeys

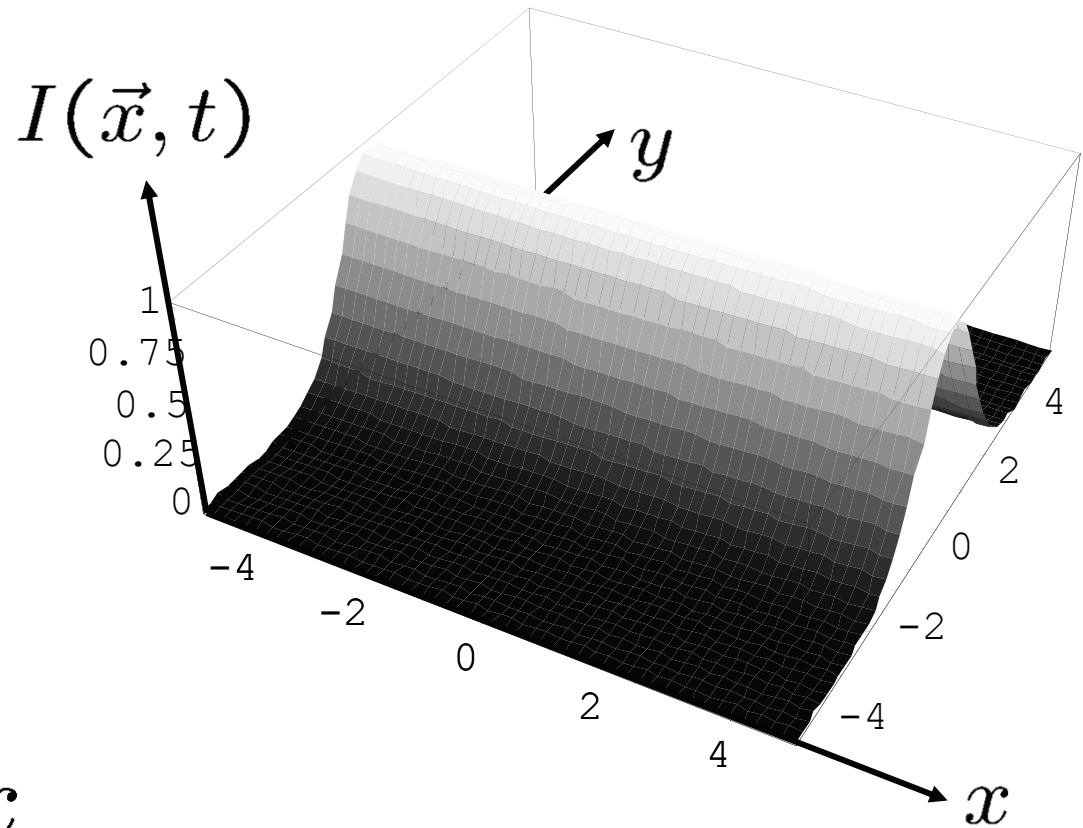
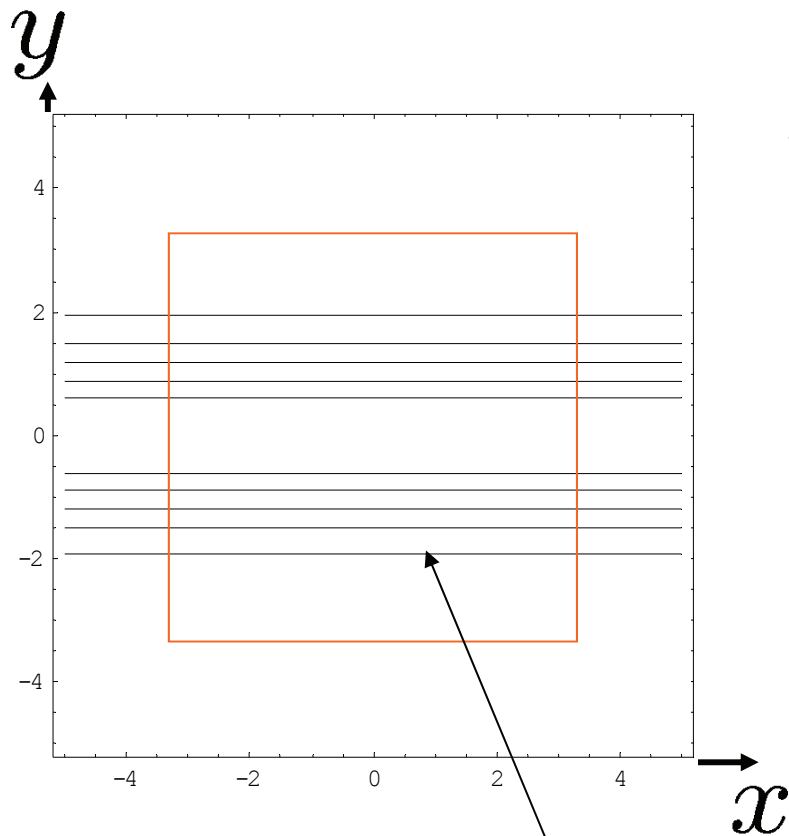
どうやらエッジの「曲がり情報が重要」

提示パターン

選択的に反応したパターン

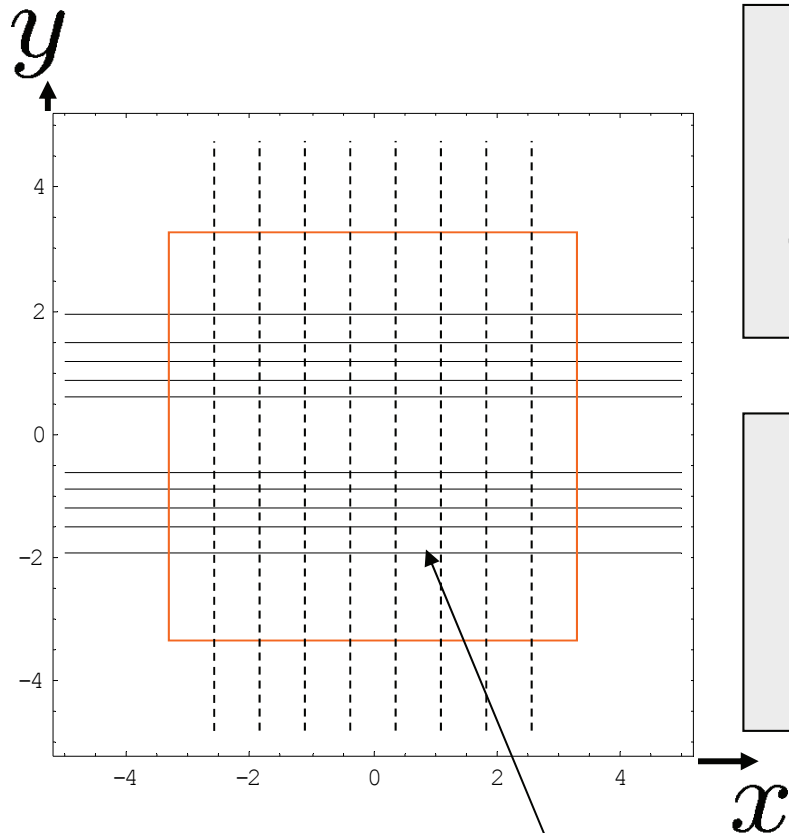


# エッジの曲がり情報 (V2情報) を定式化し 評価関数に組み込む



輝度の等高線 (レベルセット) の曲率情報を基にして考える

# エッジの曲がり情報 (V2情報) を定式化し 評価関数に組み込む



## Curvature of Level-Set

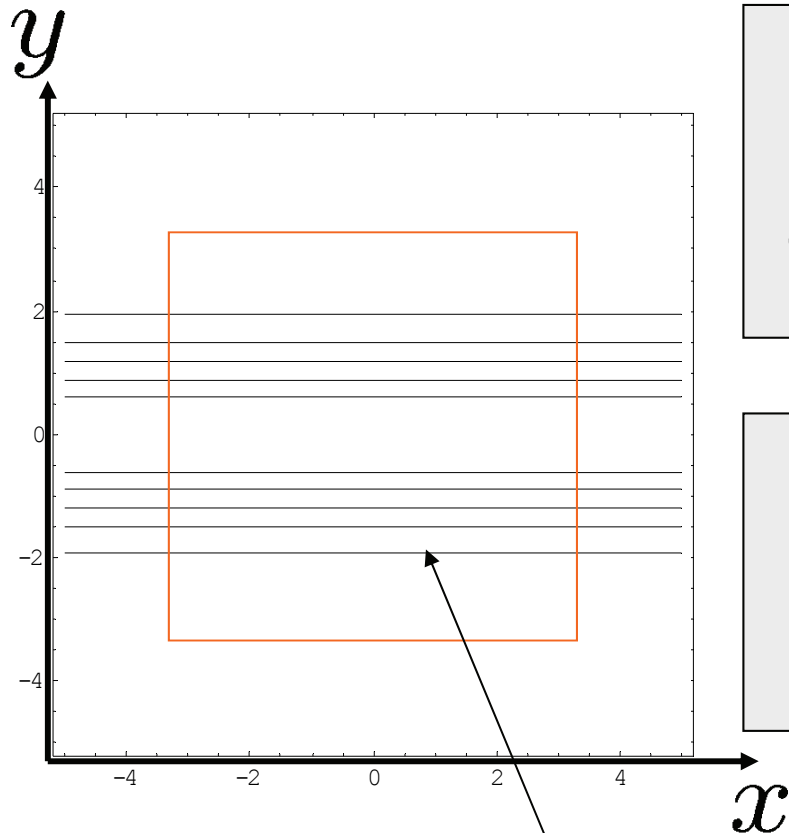
$$\kappa(\vec{x}, t) = \frac{I_x^2 I_{yy} - 2I_x I_y I_{xy} + I_y^2 I_{xx}}{(I_x^2 + I_y^2)^{3/2}}$$

## Curvature of Flow-Curve

$$\mu(\vec{x}, t) = \frac{(I_x^2 - I_y^2) I_{xy} - I_x I_y (I_{yy} - I_{xx})}{(I_x^2 + I_y^2)^{3/2}}$$

輝度の等高線 (レベルセット) の曲率情報に基づいて考える

# エッジの曲がり情報 (V2情報) を定式化し 評価関数に組み込む



## Curvature of Level-Set

$$\kappa(\vec{x}, t) = 0$$

どの位置でもレベルセットが  
曲がっていないので0

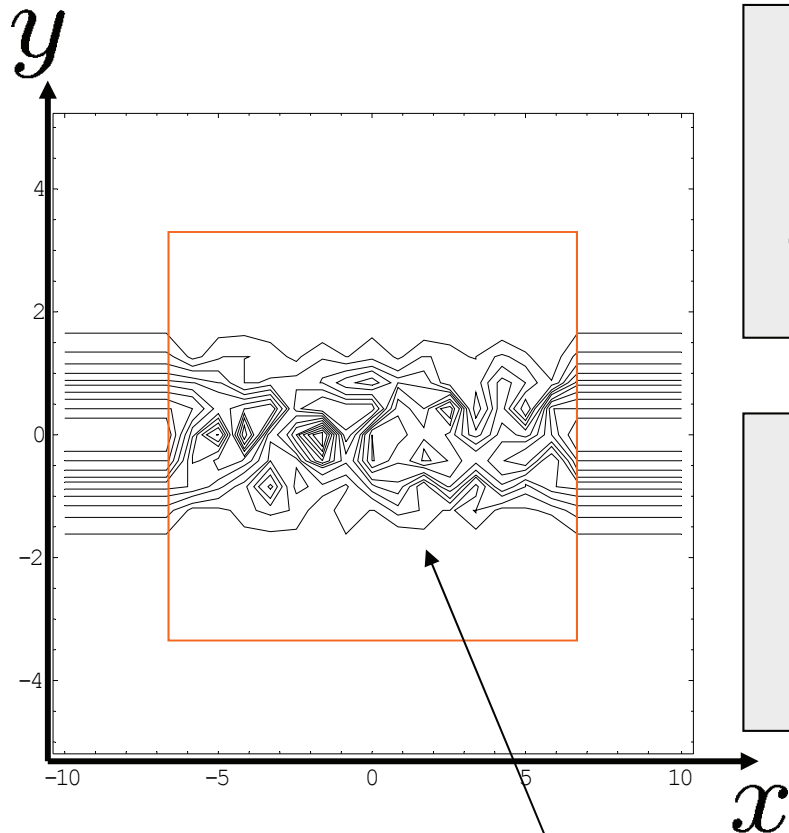
## Curvature of Flow-Curve

$$\mu(\vec{x}, t) = 0$$

どの位置でも  
隣接するレベルセットが  
「平行なので」0

輝度の等高線 (レベルセット) の曲率情報を基にして考える

# エッジの曲がり情報 (V2情報) を定式化し 評価関数に組み込む



## Curvature of Level-Set

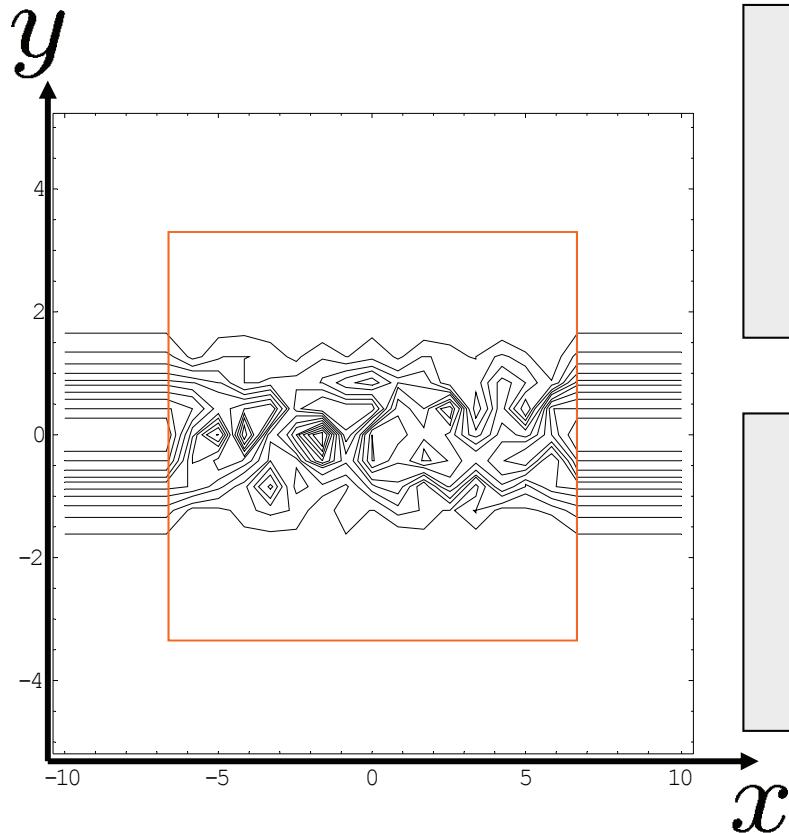
$$\kappa(\vec{x}, t) = \frac{I_x^2 I_{yy} - 2I_x I_y I_{xy} + I_y^2 I_{xx}}{(I_x^2 + I_y^2)^{3/2}}$$

## Curvature of Flow-Curve

$$\mu(\vec{x}, t) = \frac{(I_x^2 - I_y^2) I_{xy} - I_x I_y (I_{yy} - I_{xx})}{(I_x^2 + I_y^2)^{3/2}}$$

輝度の等高線 (レベルセット) の曲率情報に基づいて考える

# エッジの曲がり情報 (V2情報) を定式化し 評価関数に組み込む



## Curvature of Level-Set

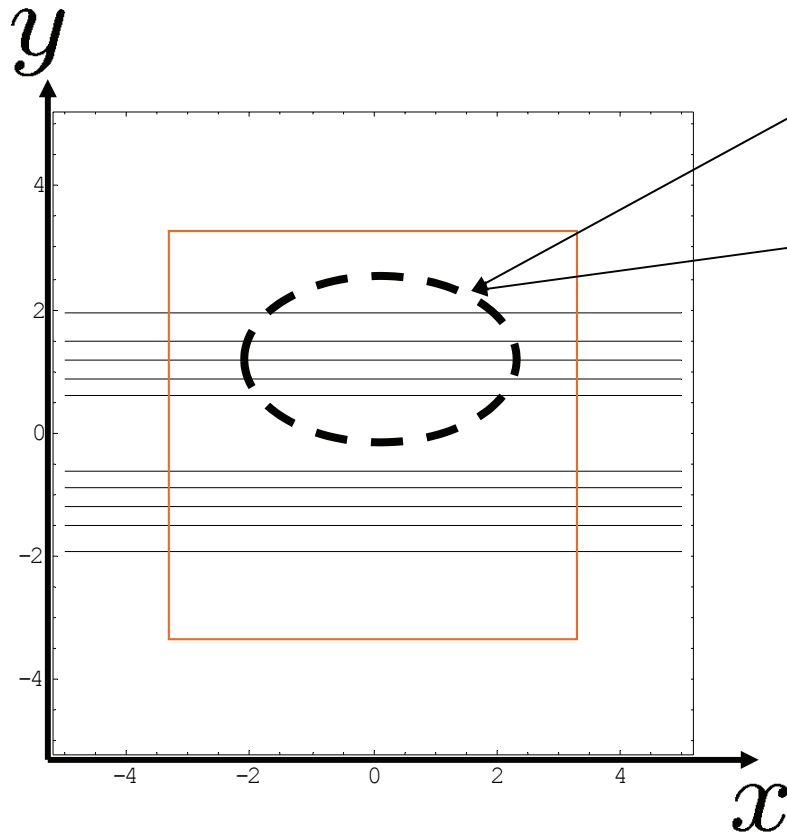
$\kappa(\vec{x}, t) \neq 0$     どの位置でもレベルセットが  
曲がっているので0ではない

## Curvature of Flow-Curve

$\mu(\vec{x}, t) \neq 0$     どの位置でも  
隣接するレベルセットが  
「平行ではないので」0ではない



# エッジの曲がり情報 (V2情報) を定式化し 評価関数に組み込む



$$\{\kappa^2(\vec{x}, t) + \mu^2(\vec{x}, t)\} = 0$$

$$I_{\xi}^2(\vec{x}, t) \neq 0$$

$$I_{\xi}^2 \{\kappa^2 + \mu^2\} \rightarrow 0$$

エッジ強度が強くても、  
グニャグニャ曲がっていなければOK

$$E[I, t] = \int_{\mathcal{B}} I_{\xi}^2(\vec{x}, t) \{\kappa^2(\vec{x}, t) + \mu^2(\vec{x}, t)\} d^2\vec{x} \rightarrow 0$$

$$E[I] = \int_{\vec{x} \in B} d^2 \vec{x} \left\{ \kappa(\vec{x}, t)^2 + \mu(\vec{x}, t)^2 \right\} I_{\xi}(\vec{x}, t)^2$$

最急降下法

$$\frac{\partial}{\partial t} I(\vec{x}; t) = ((I_x(\vec{x})^6 (I_{yyyy}(\vec{x}) I_x(\vec{x})^2 + I_{yyy}(\vec{x}) I_x(\vec{x}) I_{xy}(\vec{x}) - 9 I_{yy}(\vec{x}) I_x(\vec{x}) I_{xyy}(\vec{x}) - 4 I_{yy}(\vec{x})^2 I_{xx}(\vec{x}) + 4 I_{xy}(\vec{x})^2 I_{xx}(\vec{x}) + 2 I_x(\vec{x}) I_{xyy}(\vec{x}) I_{xx}(\vec{x}) + 5 I_x(\vec{x}) I_{xy}(\vec{x}) I_{xxy}(\vec{x}) + I_x(\vec{x})^2 I_{xxyy}(\vec{x}) + I_{yy}(\vec{x}) I_x(\vec{x}) I_{xxx}(\vec{x})) + I_y(\vec{x})^7 (I_{xx}(\vec{x}) (I_{yyy}(\vec{x}) - 9 I_{xxy}(\vec{x})) + 2 I_{yy}(\vec{x}) I_{xxy}(\vec{x}) + I_{xy}(\vec{x}) (5 I_{xyy}(\vec{x}) + I_{xxx}(\vec{x})) + 2 I_x(\vec{x}) (I_{yyy}(\vec{x}) - 3 I_{xxy}(\vec{x}))) +$$

神経生理学的妥当性はほぼ望めない

まだ何かが足りない

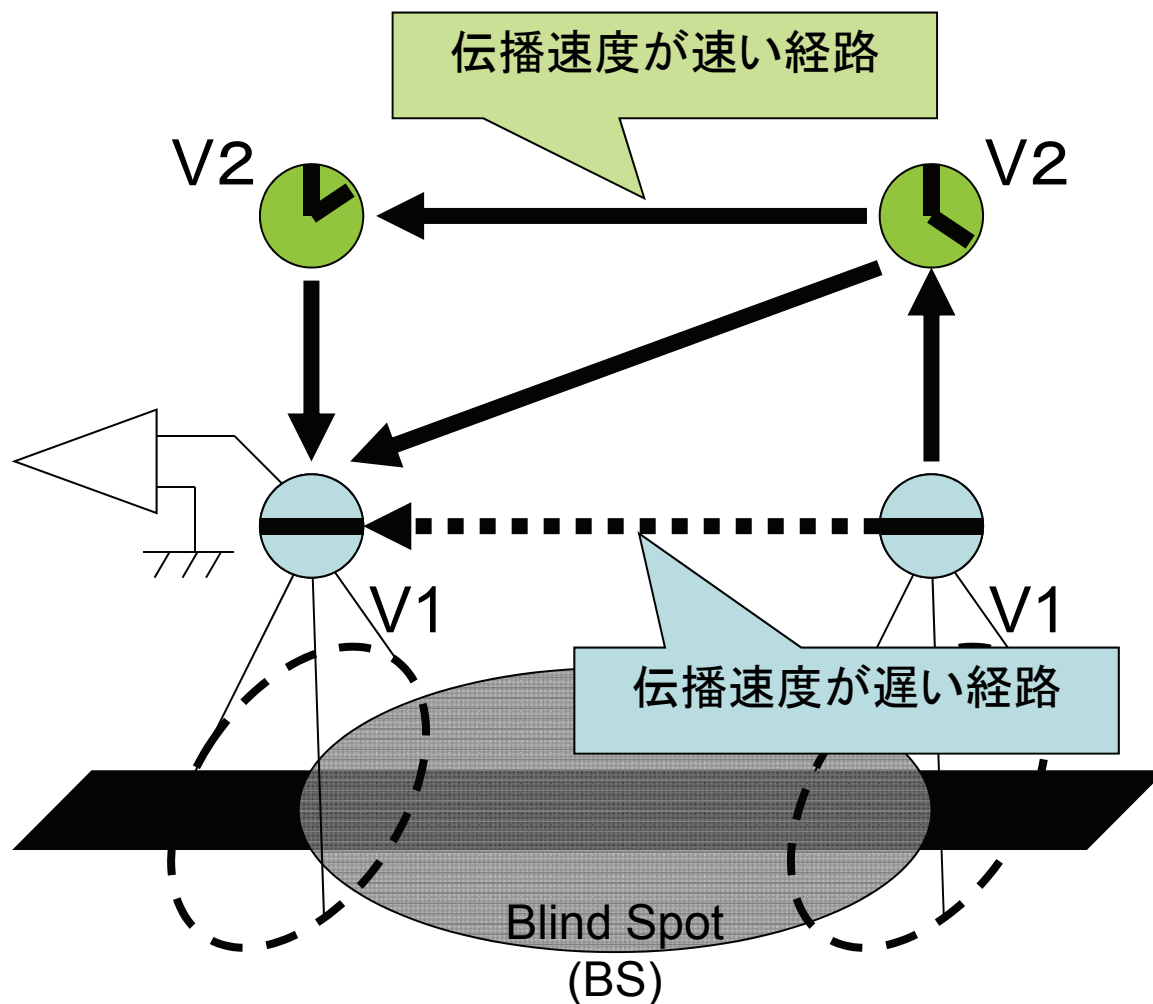
$$I_y(\vec{x})^4 I_x(\vec{x})^2 (3 I_{yyyy}(\vec{x}) I_x(\vec{x})^2 + 19 I_{yyy}(\vec{x}) I_x(\vec{x}) I_{xy}(\vec{x}) + 24 I_{yy}(\vec{x}) I_{xy}(\vec{x})^2 + 21 I_{yy}(\vec{x}) I_x(\vec{x}) I_{xyy}(\vec{x}) + 12 I_{yy}(\vec{x})^2 I_{xx}(\vec{x}) - 12 I_{xy}(\vec{x})^2 I_{xx}(\vec{x}) + 6 I_x(\vec{x}) I_{xyy}(\vec{x}) I_{xx}(\vec{x}) - 24 I_{yy}(\vec{x}) I_{xx}(\vec{x})^2 - 33 I_x(\vec{x}) I_{xy}(\vec{x}) I_{xxy}(\vec{x}) + 6 I_x(\vec{x})^2 I_{xxyy}(\vec{x}) - 13 I_{yy}(\vec{x}) I_x(\vec{x}) I_{xxx}(\vec{x}) + 3 I_x(\vec{x})^2 I_{xxxx}(\vec{x})) +$$

$$I_y(\vec{x})^6 (I_{yyyy}(\vec{x}) I_x(\vec{x})^2 + 5 I_{yyy}(\vec{x}) I_x(\vec{x}) I_{xy}(\vec{x}) + 4 I_{yy}(\vec{x}) I_{xy}(\vec{x})^2 + 3 I_{yy}(\vec{x}) I_x(\vec{x}) I_{xyy}(\vec{x}) + 2 I_x(\vec{x}) I_{xyy}(\vec{x}) I_{xx}(\vec{x}) - 4 I_{yy}(\vec{x}) I_{xx}(\vec{x})^2 - 7 I_x(\vec{x}) I_{xy}(\vec{x}) I_{xxy}(\vec{x}) + 4 I_x(\vec{x})^2 I_{xxyy}(\vec{x}) - 3 I_{yy}(\vec{x}) I_x(\vec{x}) I_{xxx}(\vec{x}) + 3 I_x(\vec{x})^2 I_{xxxx}(\vec{x}))) / (I_y(\vec{x})^2 + I_x(\vec{x})^2)^3$$

$$E[I] = \int_{\vec{x} \in B} d^2\vec{x} \left\{ \kappa(\vec{x}, t)^2 + \mu(\vec{x}, t)^2 \right\} I_{\xi}(\vec{x}, t)^2$$

V2細胞がコードしている情報

V1細胞がコードしている情報



先に最適化

$$\kappa(\vec{x}, t)$$

$$\mu(\vec{x}, t)$$

更新速い

後で最適化

$$I_{\xi}(\vec{x}, t)$$

↓

$$I_{\xi}(\vec{x})$$

更新遅い

$$E[I] = \int_{\vec{x} \in B} d^2 \vec{x} \left\{ \kappa(\vec{x}, t)^2 + \mu(\vec{x}, t)^2 \right\} I_{\xi}(\vec{x}, t)^2$$

最急降下法

時間に依存しない量と仮定

$$\frac{\partial}{\partial t} I(\vec{x}, t) = \frac{\partial}{\partial \eta} \kappa(\vec{x}, t) I_{\xi}(\vec{x}, t)^2 + \frac{\partial}{\partial \xi} \mu(\vec{x}, t) I_{\xi}(\vec{x}, t)^2 - \Delta \kappa(\vec{x}, t) I_{\xi}^2(\vec{x}, t)$$

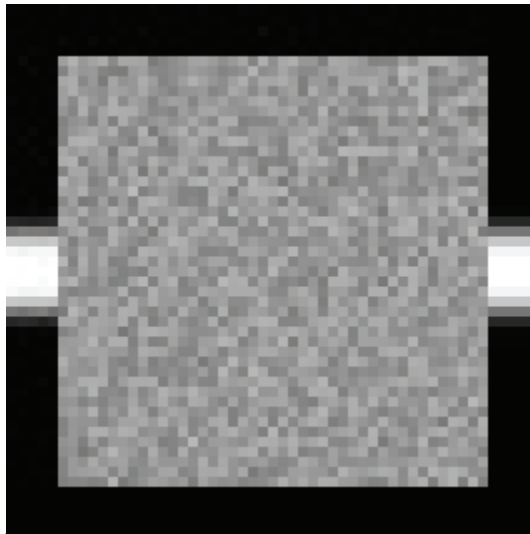
実にすっきり！

しかし、本当にこの式で望みの解が得られるのだろうか？

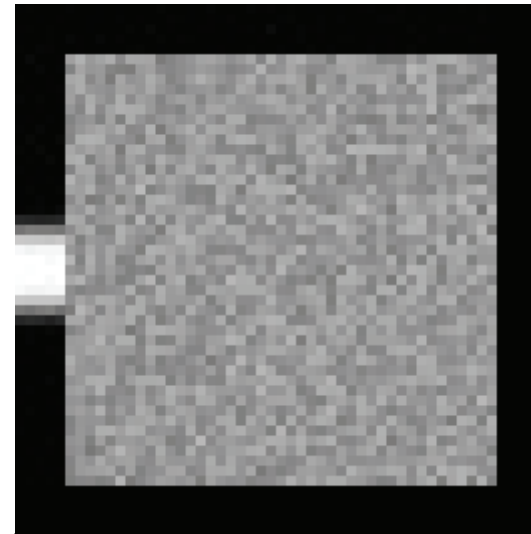
## 数値シミュレーションで確認する

$$\frac{\partial}{\partial t} I(\vec{x}, t) = \frac{\partial}{\partial \eta} \kappa(\vec{x}, t) I_{\xi}(\vec{x}, t)^2 + \frac{\partial}{\partial \xi} \mu(\vec{x}, t) I_{\xi}(\vec{x}, t)^2 - \Delta \kappa(\vec{x}, t) I_{\xi}^2(\vec{x}, t)$$

## 結果



BSの両端に  
バーがある場合は、  
充填して一本の  
長いバーに

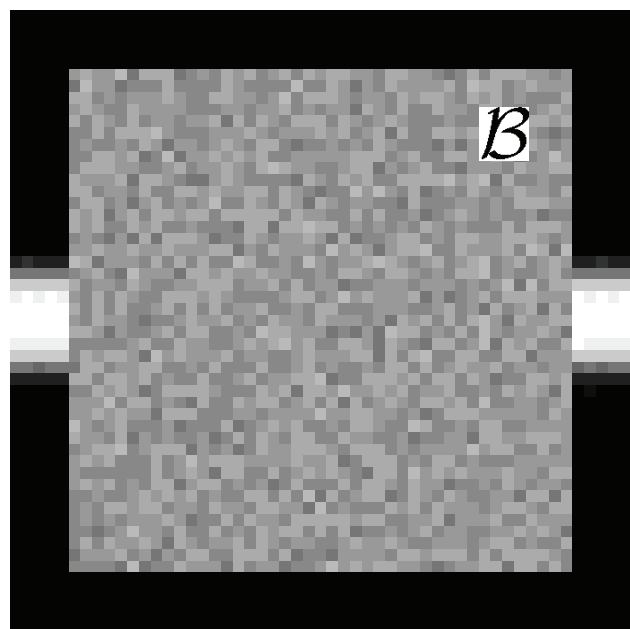


BSの片側にしか  
バーがない場合は、  
充填されず  
短いバーのまま

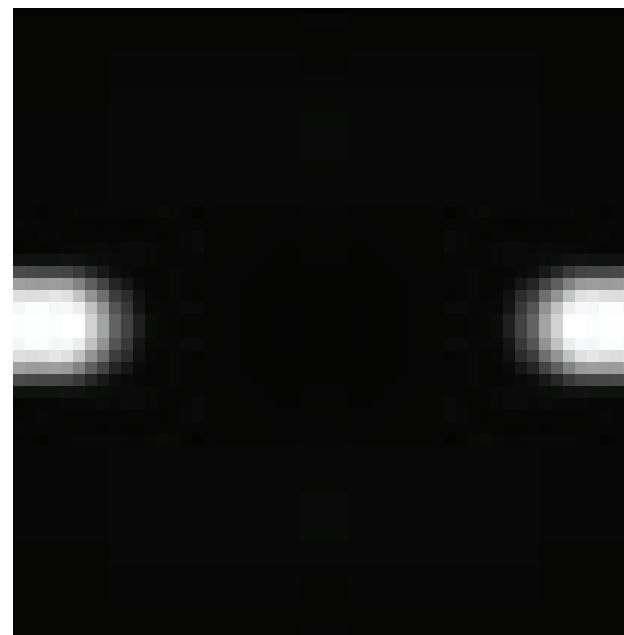
# V2細胞の情報があればうまくいくが

$$E[I] = \int_{\vec{x} \in B} d^2\vec{x} \left\{ \kappa(\vec{x}, t)^2 + \mu(\vec{x}, t)^2 \right\} I_{\xi}^2(\vec{x}, t)$$

V1                  V2                                  V1



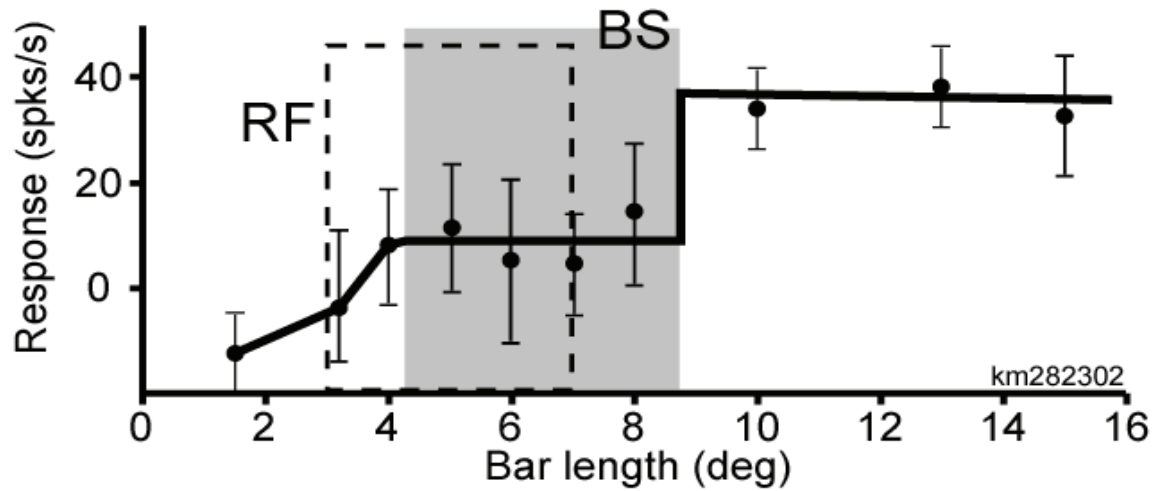
$I(\vec{x}, t = 0)$



$I^*(\vec{x})$  現模

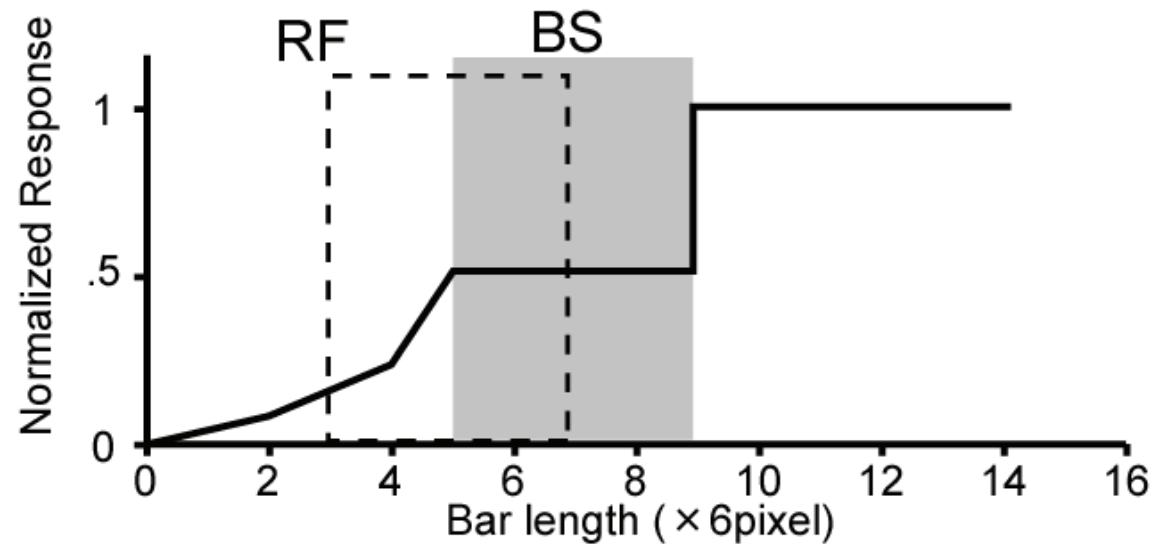
# 神経生理学的妥当性

M. Matsumoto & H. Komatsu,  
*J Neurophysiol.* (2005)



生理データ

モデル

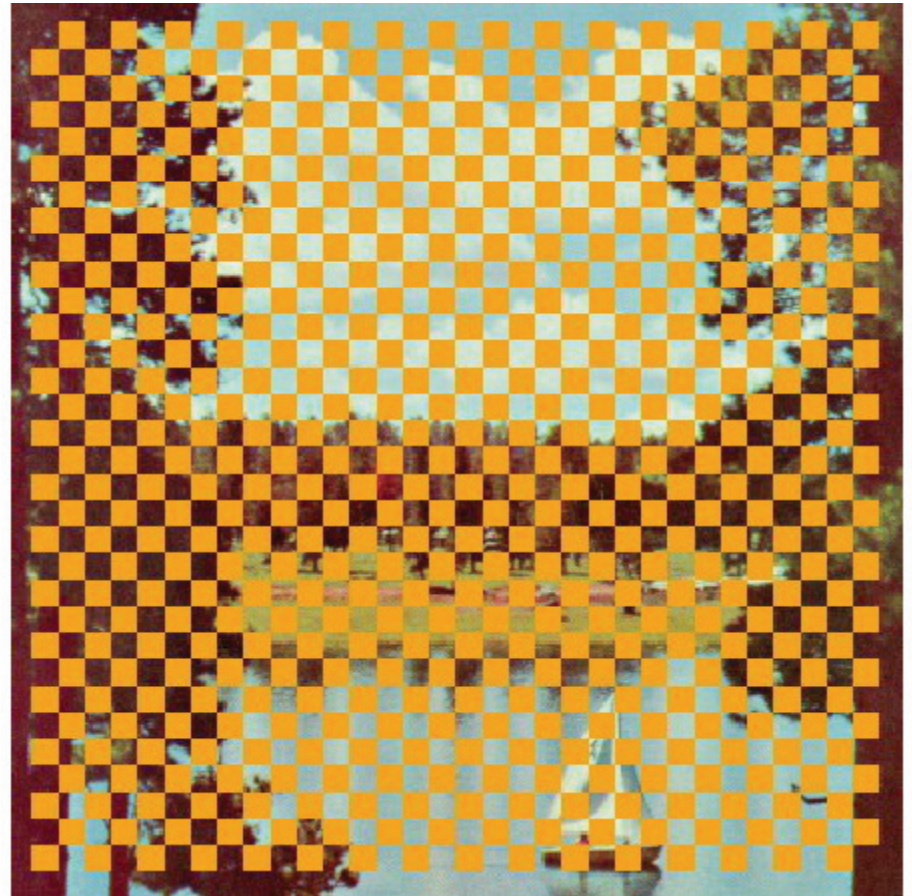
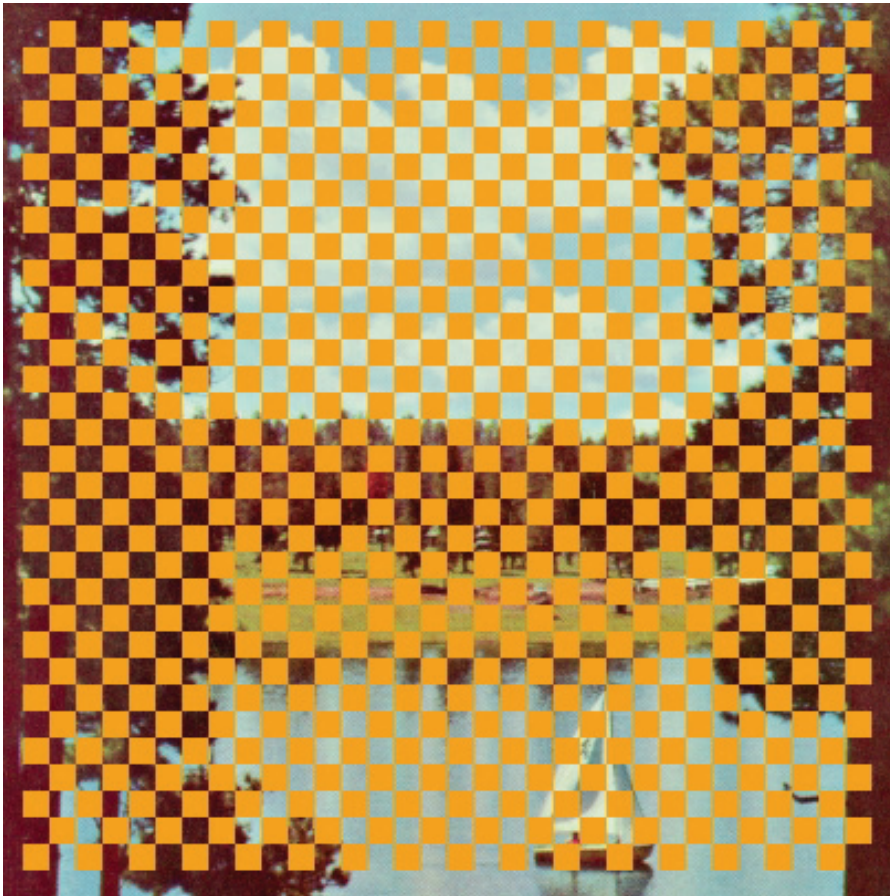


# Image Inpainting としての有効性

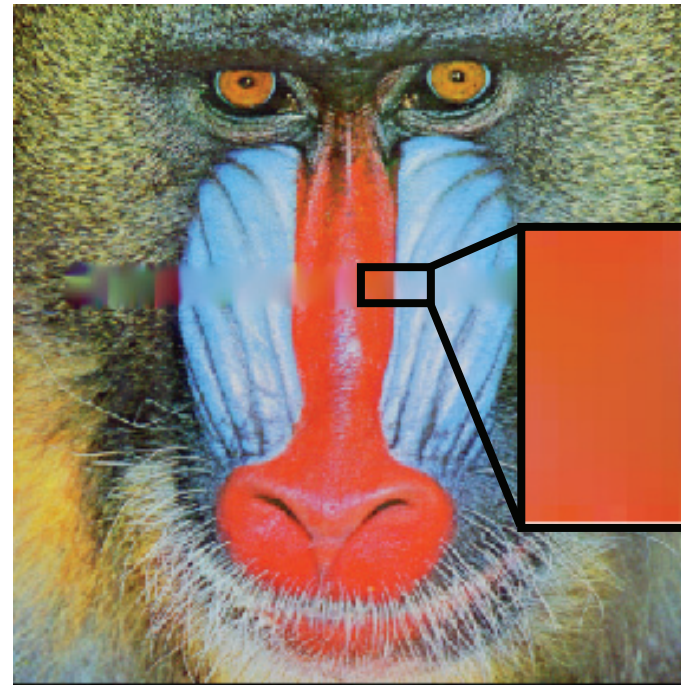
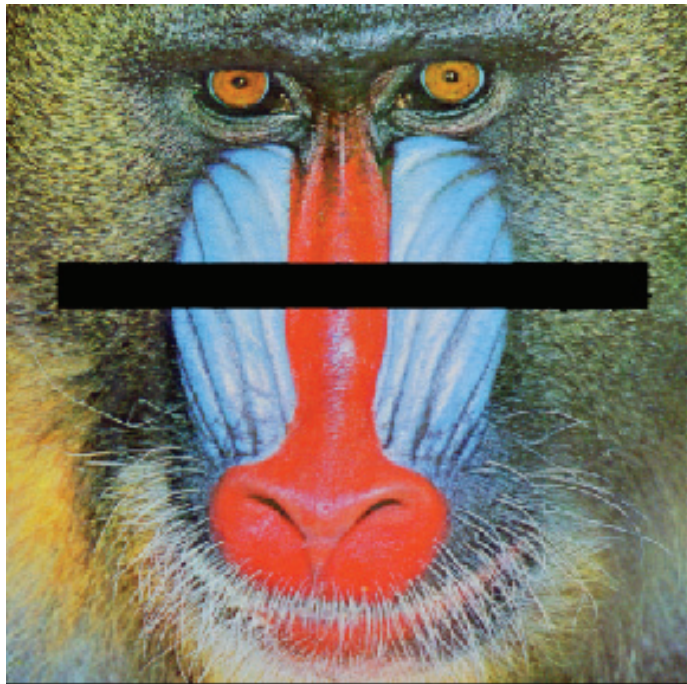




# 極端な例

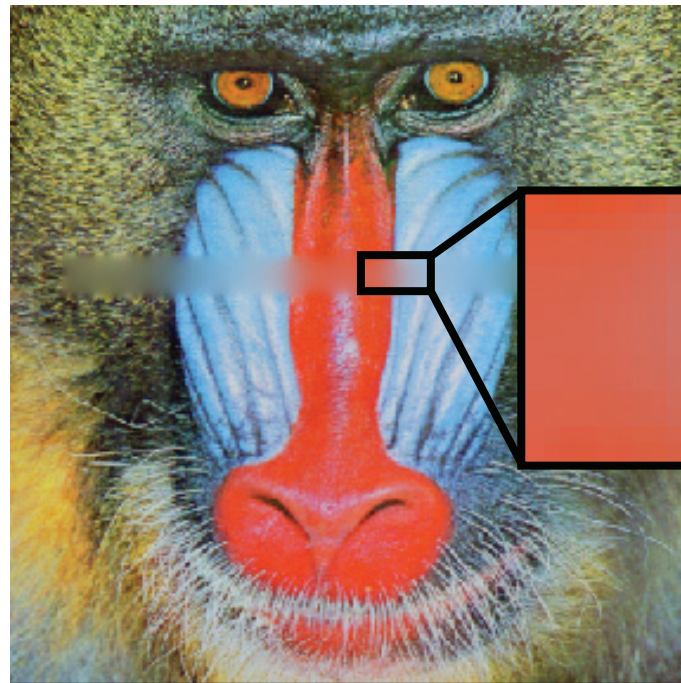


視覚モデルによる充填



Adobe Photoshop

スポット修正ツール



形状情報も  
重要だから  
(画像の事前知識)

## 考察・まとめ

ダイナミクスが  
簡単になるし、  
望みの解も得られるから

なぜV2細胞が  
必要なのか？

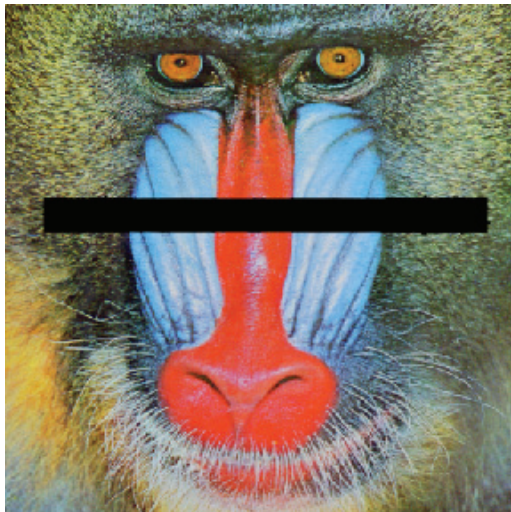
なぜ速度が異なる  
経路が必要なのか？

盲点

盲点充填を  
視覚数理モデル

なぜ？  
(今後の課題)

工学的有効性は？  
(既存手法との比較)

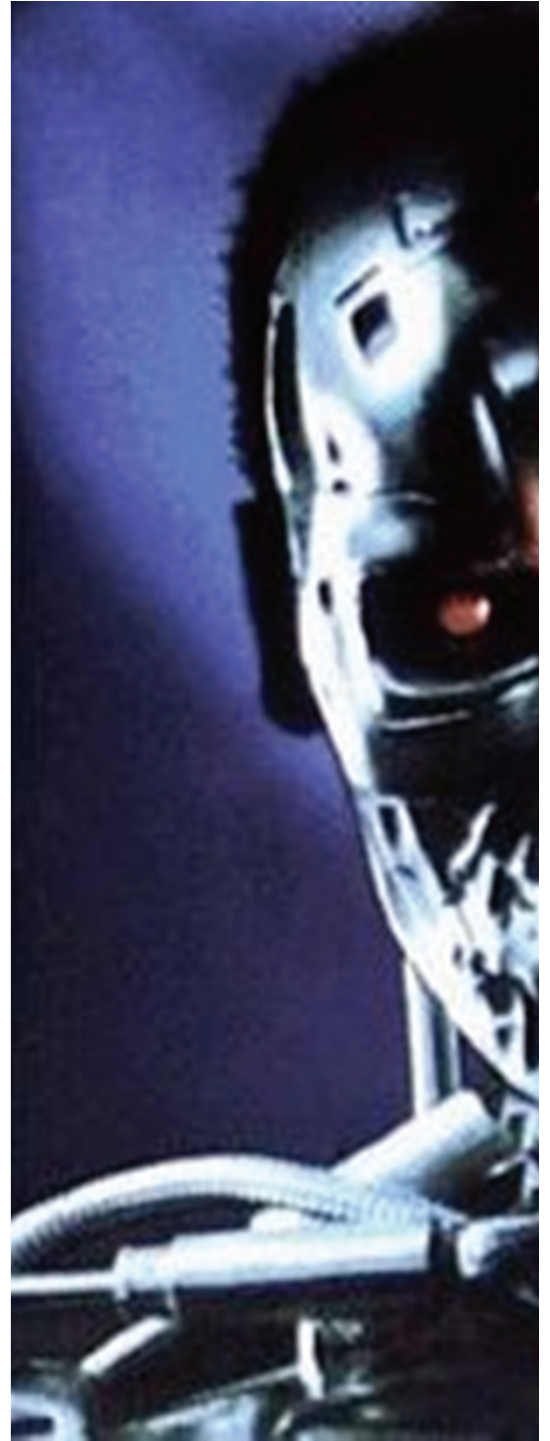




視覚  
||

画像処理マシン

と割り切った考え



# 研究手段・方針

- 視覚＝画像処理 と割り切って考える。
  - 視覚は,
    - すでに様々な機能がインストールされている画像処理機械
    - ただし, 工学的には考えられない妙な性質あり
  - 視覚を理解するためには,
    - 画像処理ハンドブックを2冊買い, 研究室と自宅に置く.
    - 時間を見つけて読んでおく.
    - 妙な性質は「単なる不具合」によるものなのか, 「必然的に生じる特性」なのかを区別する.
  - 視覚を応用するためには,
    - 数理モデルを構築し, シミュレーションする.
    - 「あえてアホなことを考える時間」を設ける(電車の中などで)
  - 便利な用語を信じない
    - コントラスト, 空間周波数, 時間周波数, 両眼視差など.