

脳が行なっているような 画像処理AIを作ってみよう

2019/03/27, 28

電気通信大学 大学院情報理工学研究科
電気通信大学 脳科学ライフサポートセンター
電気通信大学 人工知能先端研究センター

庄野 逸: shouno@uec.ac.jp

自己紹介+TA紹介

□ 氏名: 庄野 逸 (しょうの はやる)

□ 略歴:

大阪大学→奈良女子大学→山口大学
→電気通信大学

□ 興味のあること:

画像処理, 医用画像処理, ニューラルネットワーク,
脳/視覚モデル, 機械学習

□ TA 遠藤 瑛泰 (えんどう あきひろ)
庄野研 M2

□ TA 小林 源太 (こばやし げんた)
庄野研 M1



1時間目

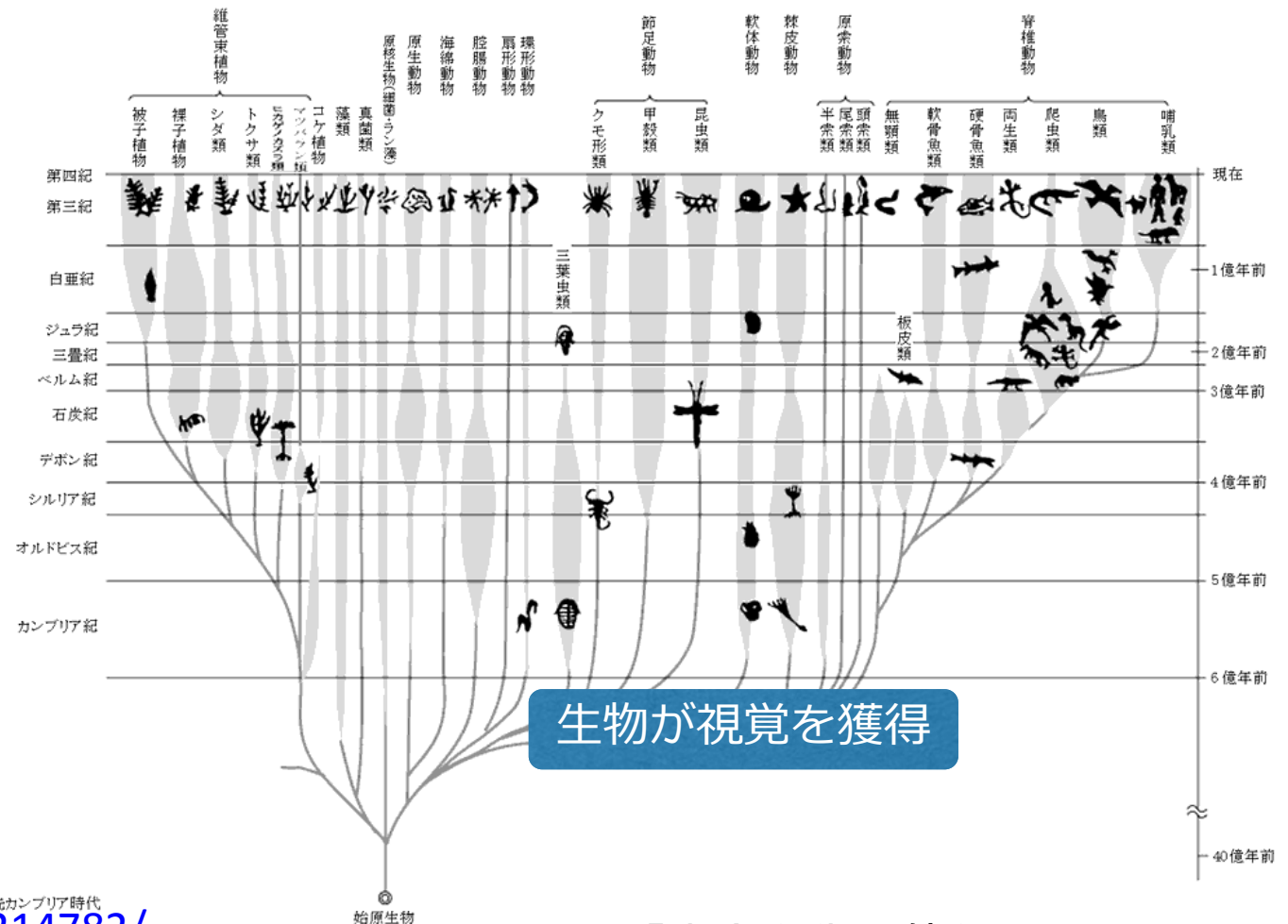
生物の視覚システム(1)

原始的な視覚システム

視覚の獲得と生存性の向上

□ カンブリア爆発

- 生物が視覚を獲得することで「種」が爆発的に増加

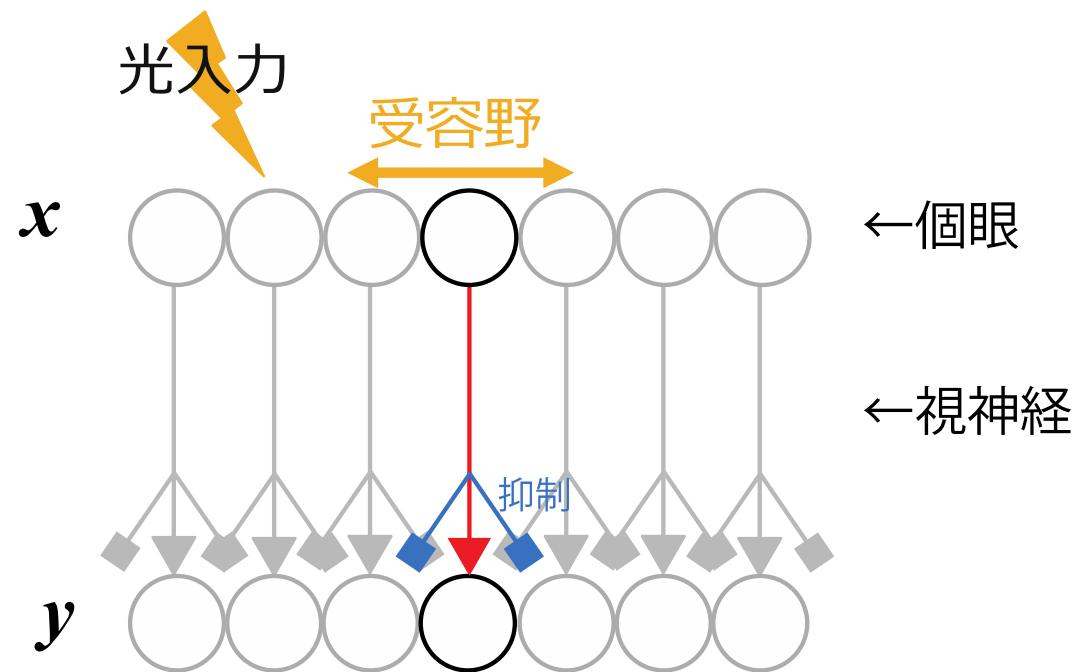


<https://www.amazon.co.jp/dp/4794214782/>

「生命進化40億年の風景」より

初期の視覚システム(1)

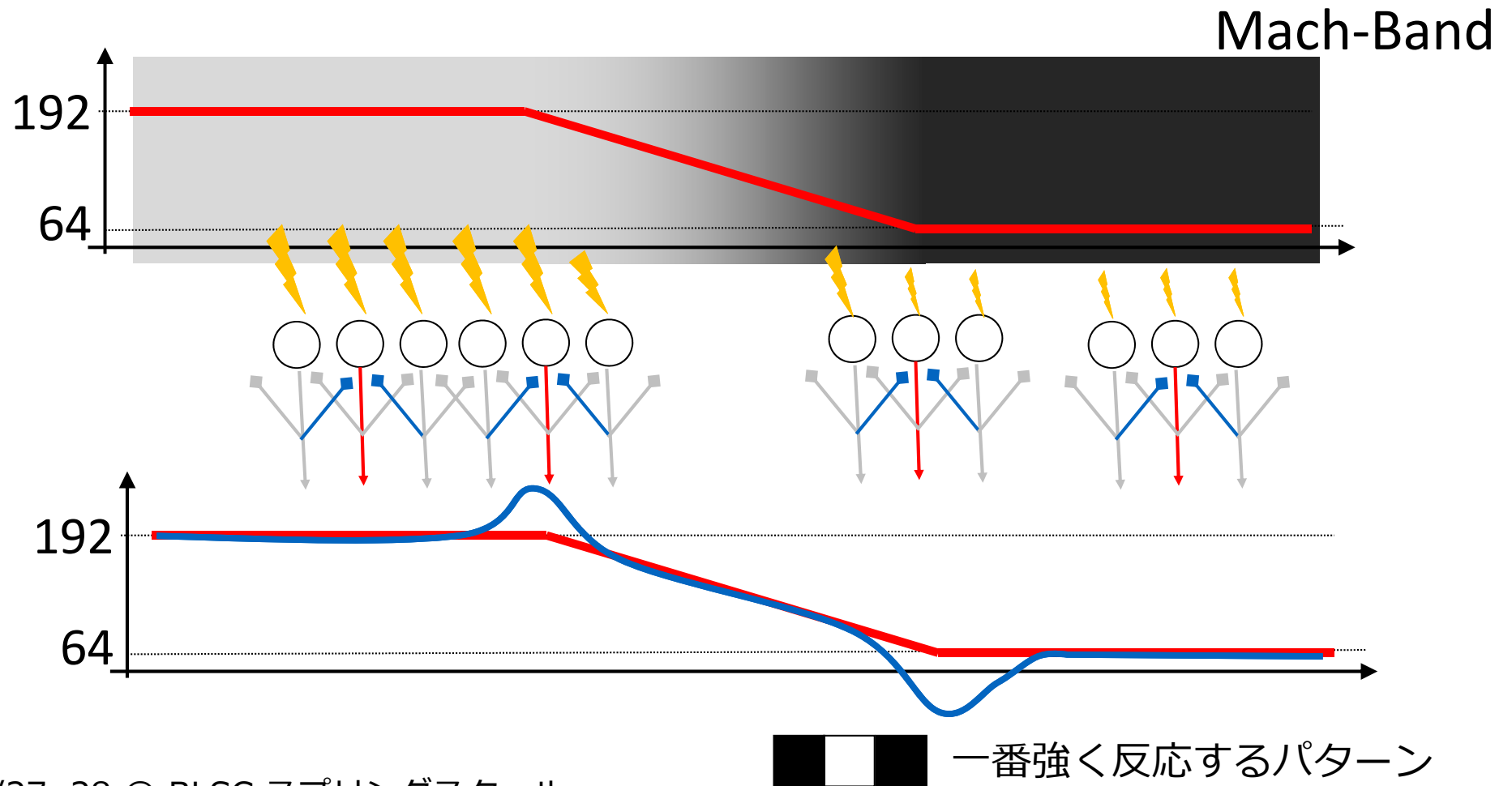
- 三葉虫は複眼システムを持っていた
- 現代ではカブトガニに近い複眼システム持っている
 - この複眼システムは、**側抑制**による信号制御を行う
[Ratliff 65]

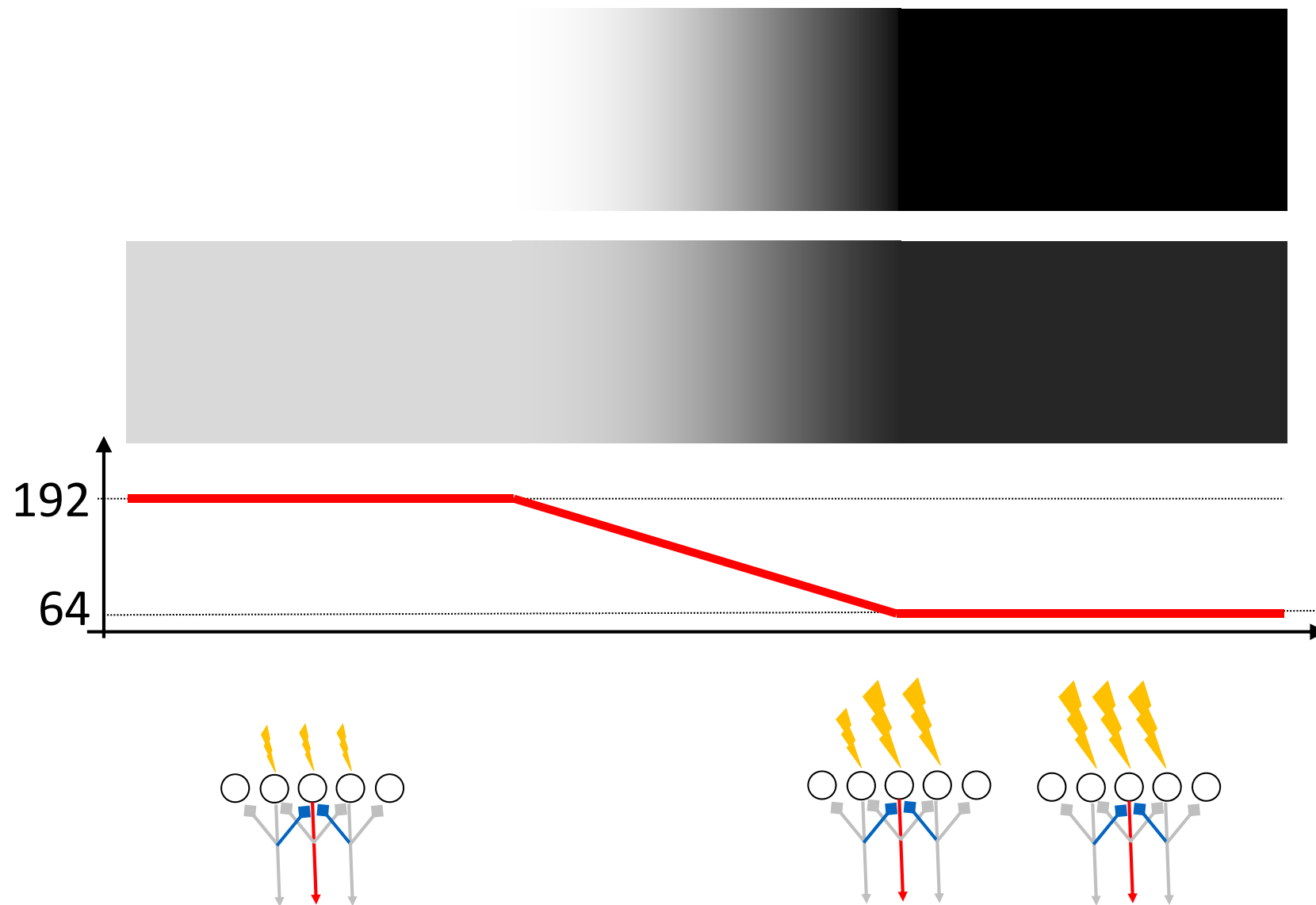


自分の周りの細胞の信号を抑制
空間フィルタと呼ばれる演算

初期の視覚システム(2)

- 三葉虫は複眼システムを持っていた
- 現代ではカブトガニに近い複眼システム持っている
- この複眼システムは、**側抑制**による信号制御を行う



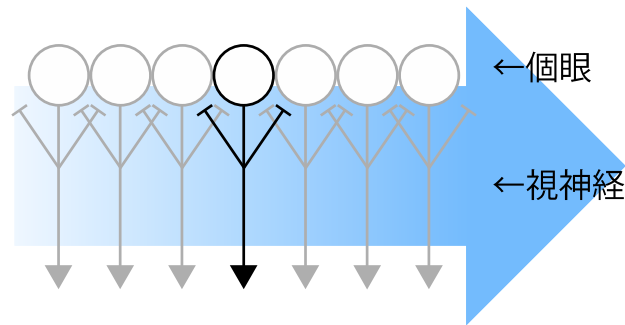


2時間目

画像のフィルタリング

初期の視覚システムとフィルタリング

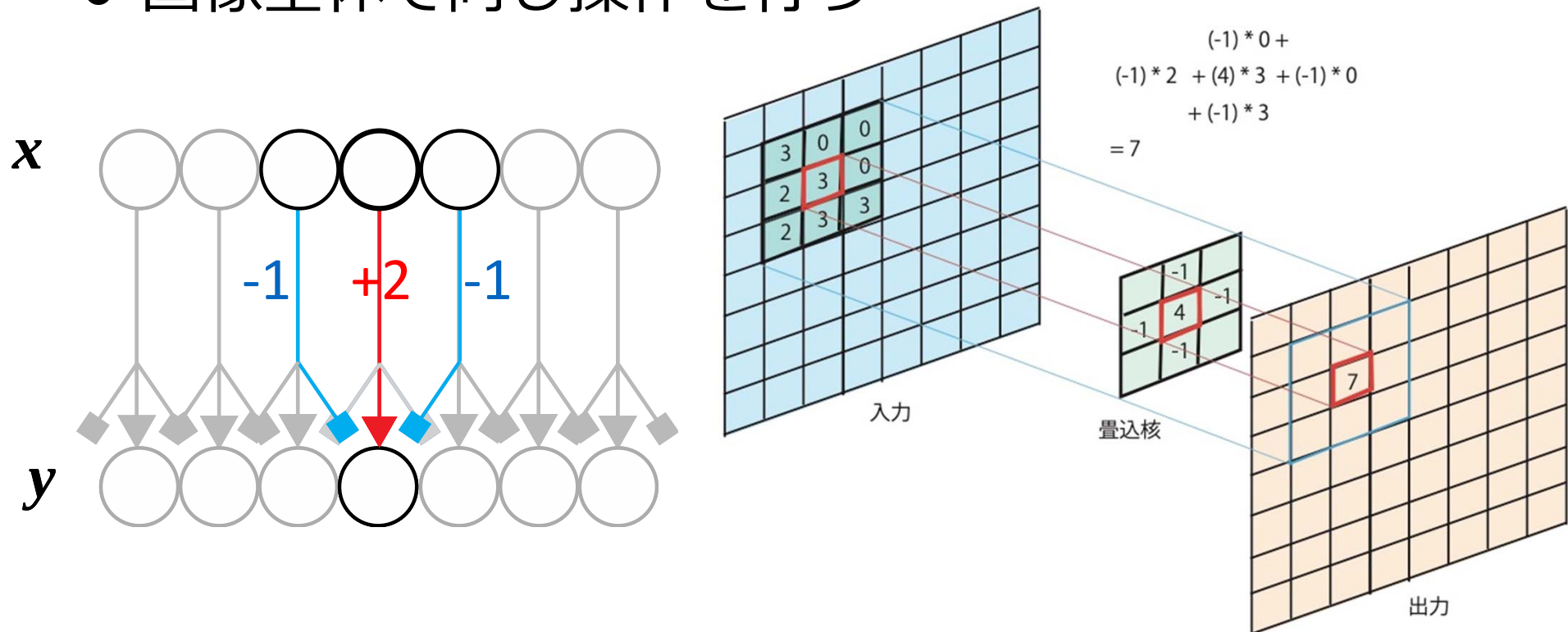
- 三葉虫は複眼システムを持っていた
- 現代ではカブトガニに近い複眼システム持っている
- この複眼システムは、**側抑制**による信号制御を行う



- 側抑制は画像のエッジ情報に強く反応する
- 単純なフィルタ結果でも結構役に立っていた

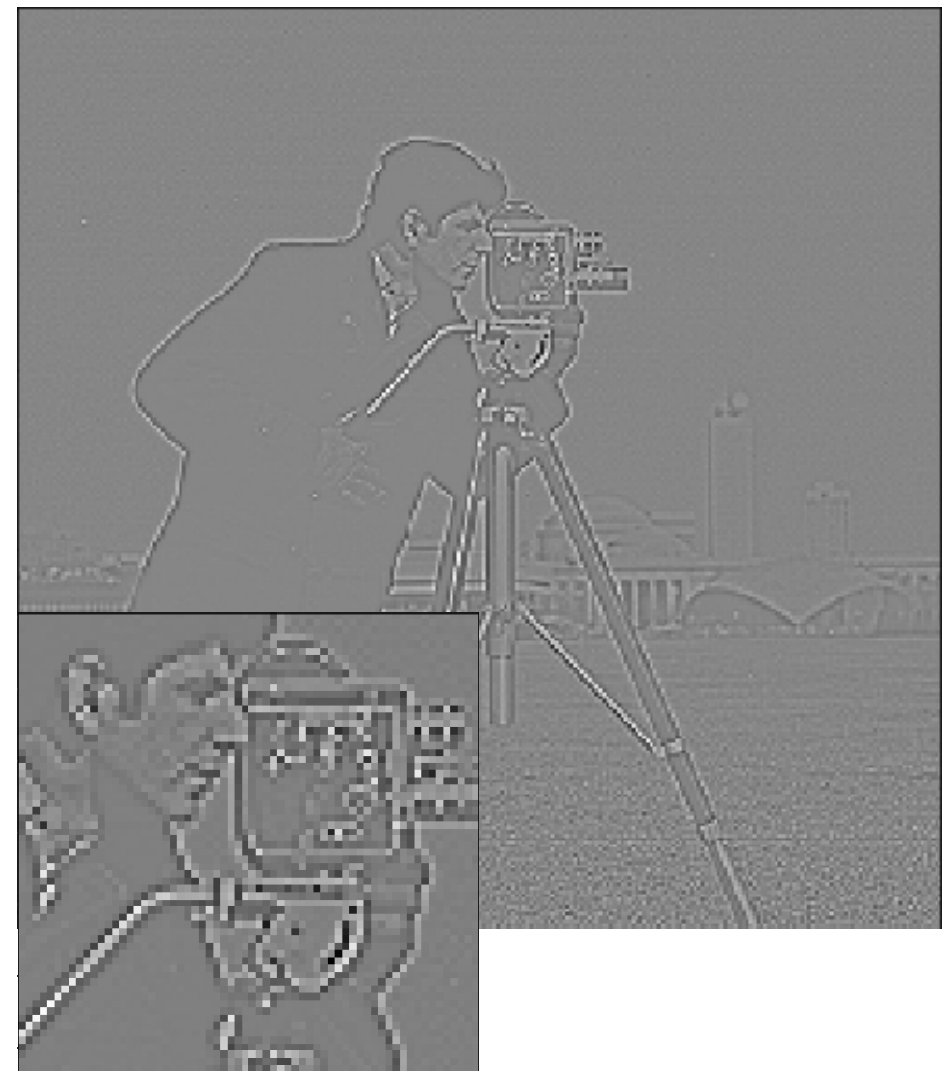
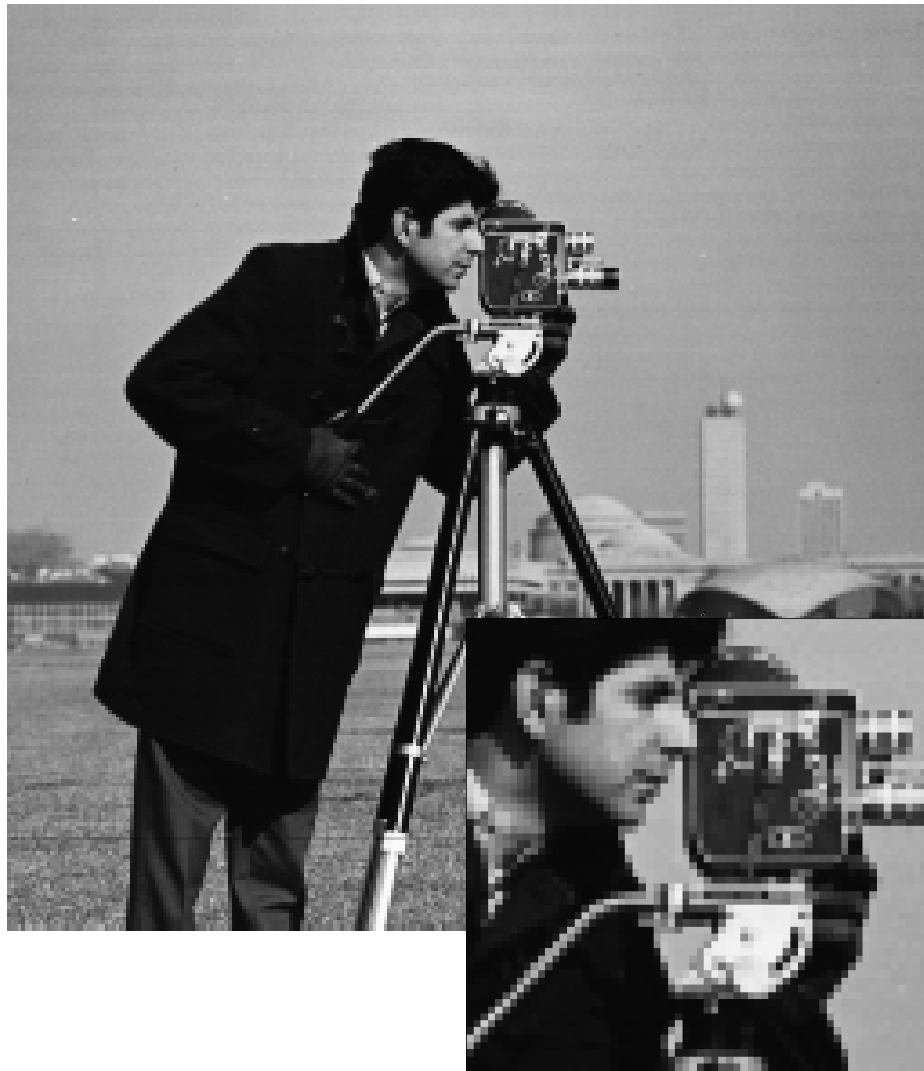
フィルタリング演算

- 入力を画像の場合は2次元の構造を持つ
- フィルタ演算は
 - ある画素の周辺画素に対して係数をかけて足し込む
側抑制は周辺が抑制で中心が興奮
 - 画像全体で同じ操作を行う



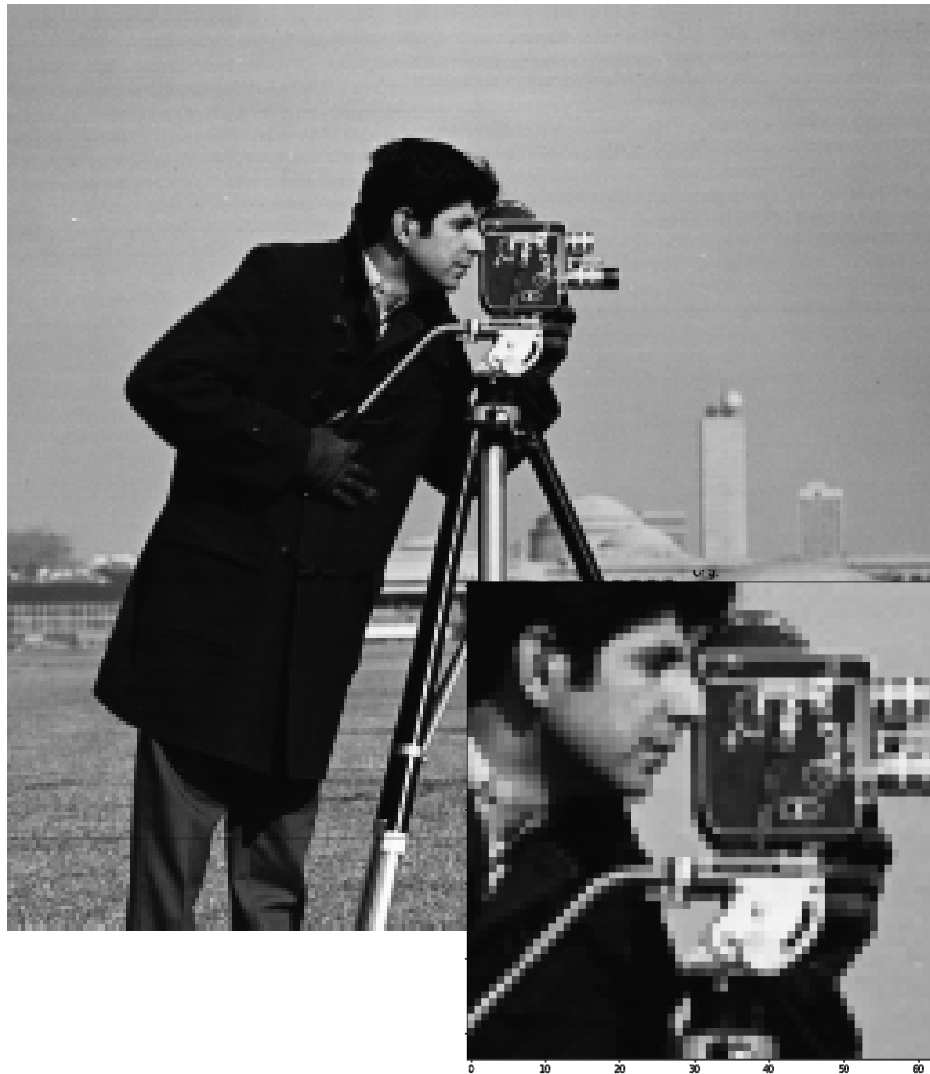
微分型のフィルタの適用

- 微分型は明るさが変化する場所（エッジ）を検出

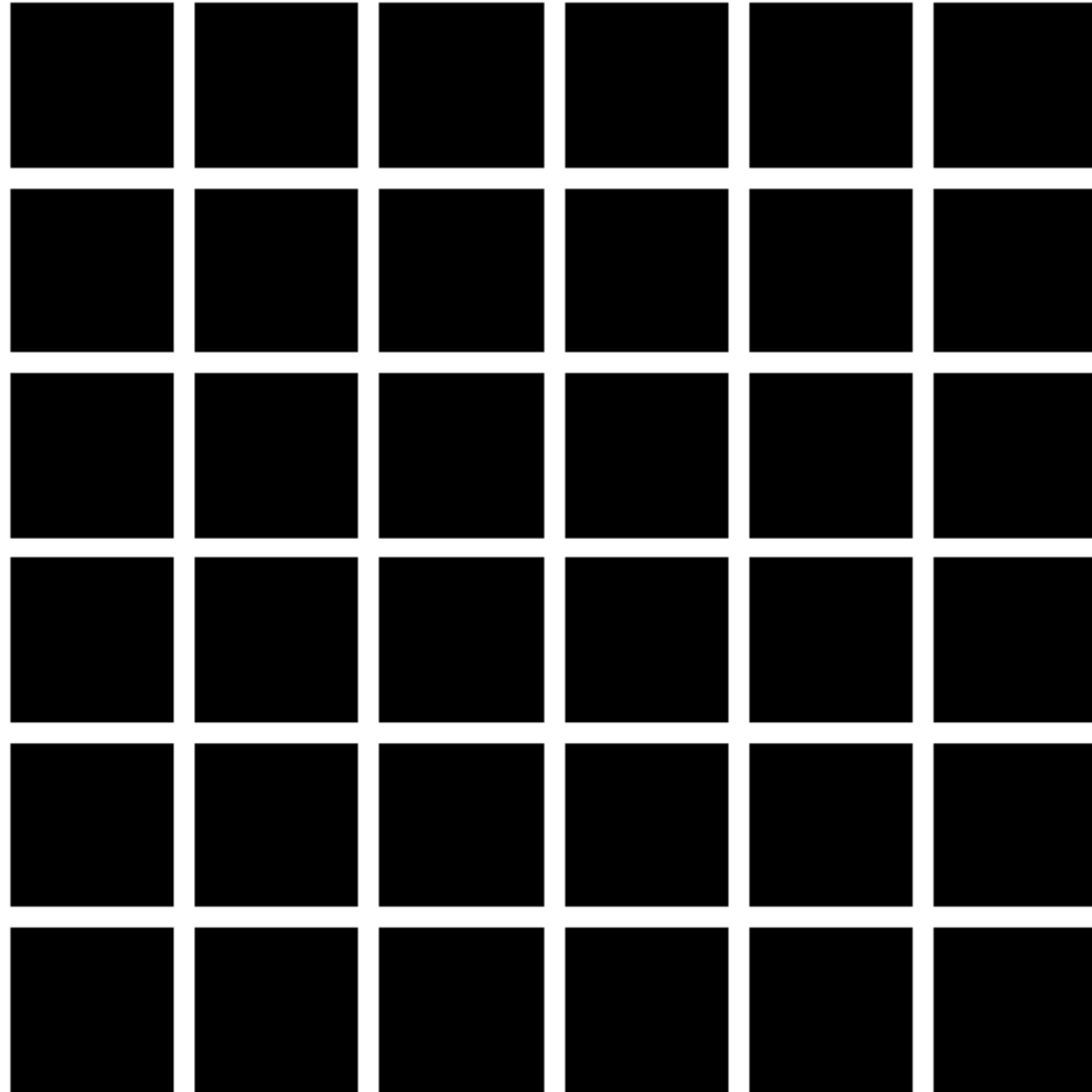


積分型のフィルタの適用

- 局所積分型は画像をぼかす作用

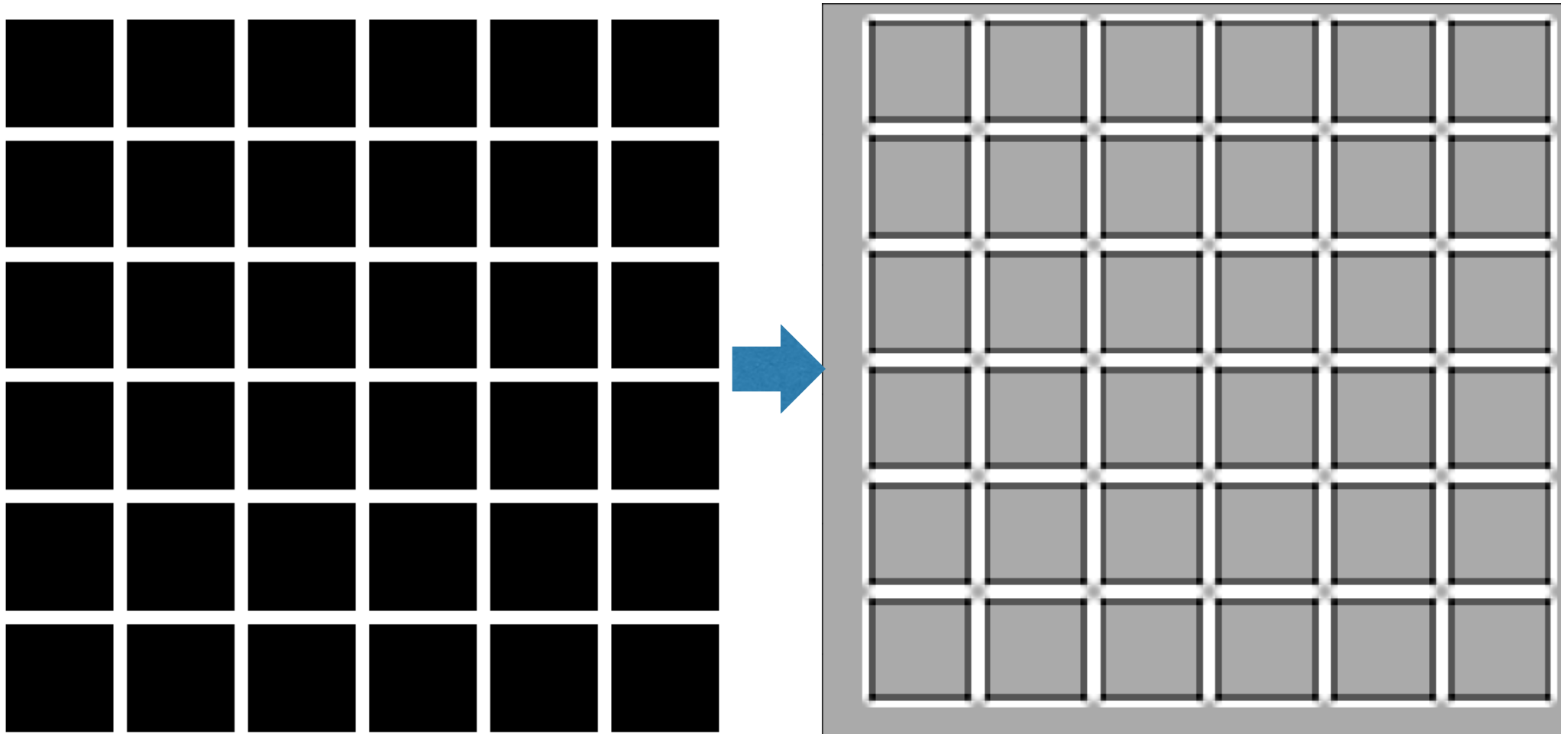


Hermann 錯視



Hermann 錯視はフィルタで説明できるか？

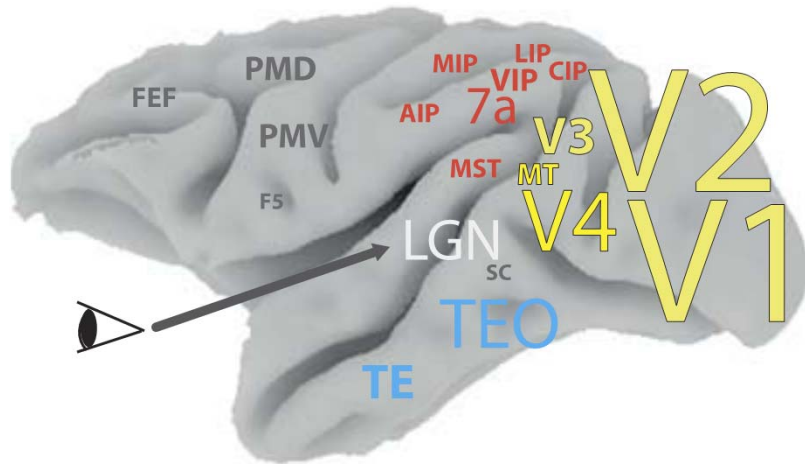
- 画像に微分形フィルタと積分形フィルタを適用



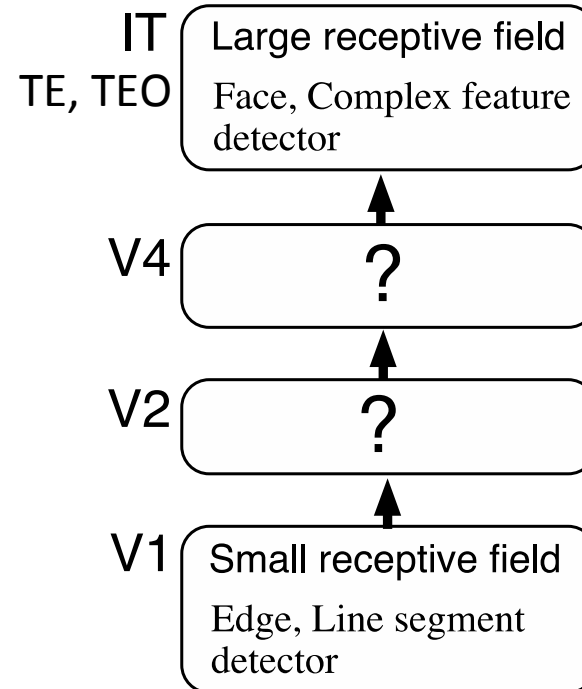
交差点部分に明るさが異なる成分が発現する

生物の視覚システム ヒトやサルの視覚システム

視覚情報の処理過程



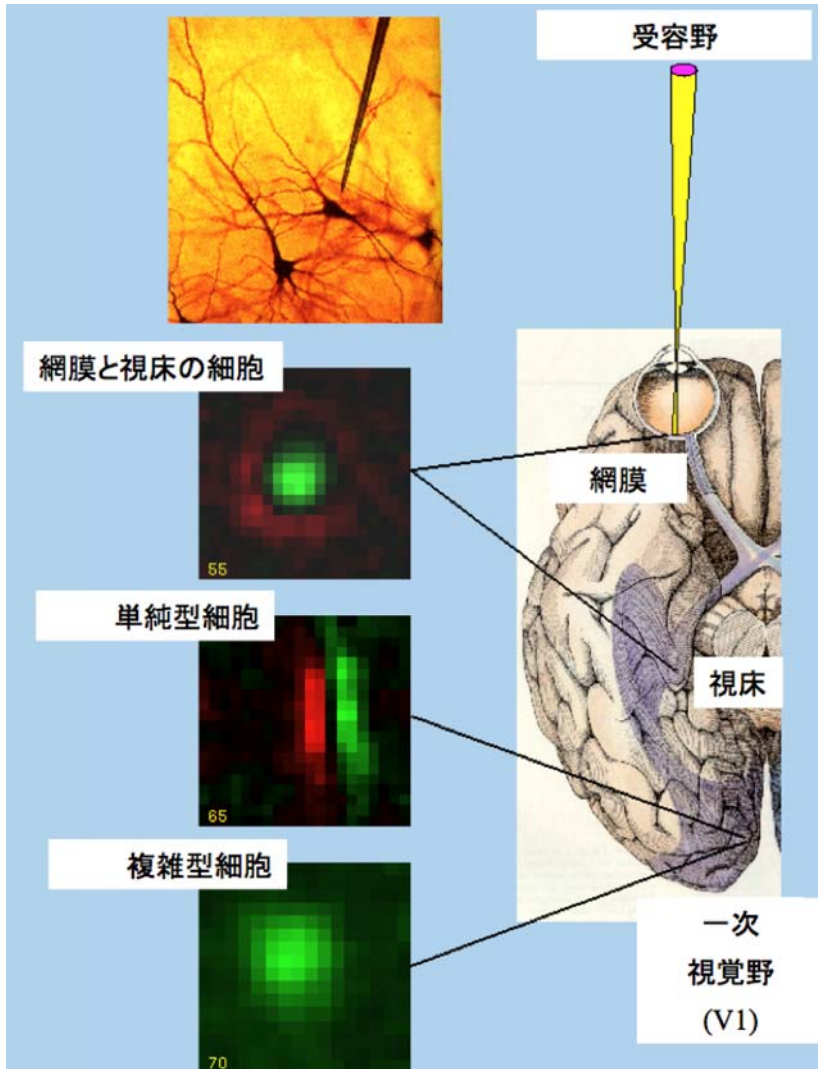
[Felleman+91, DiCarlo+12, Kruger+13]



- 視覚野: 階層構造を持ち, 階層ごとに視覚課題を解決
- 初期視覚野: 狭い受容野, 単純な特徴抽出
単純型細胞, 複雑型細胞の存在
- 高次視覚野: 広い受容野, 中程度に複雑な特徴に選択的

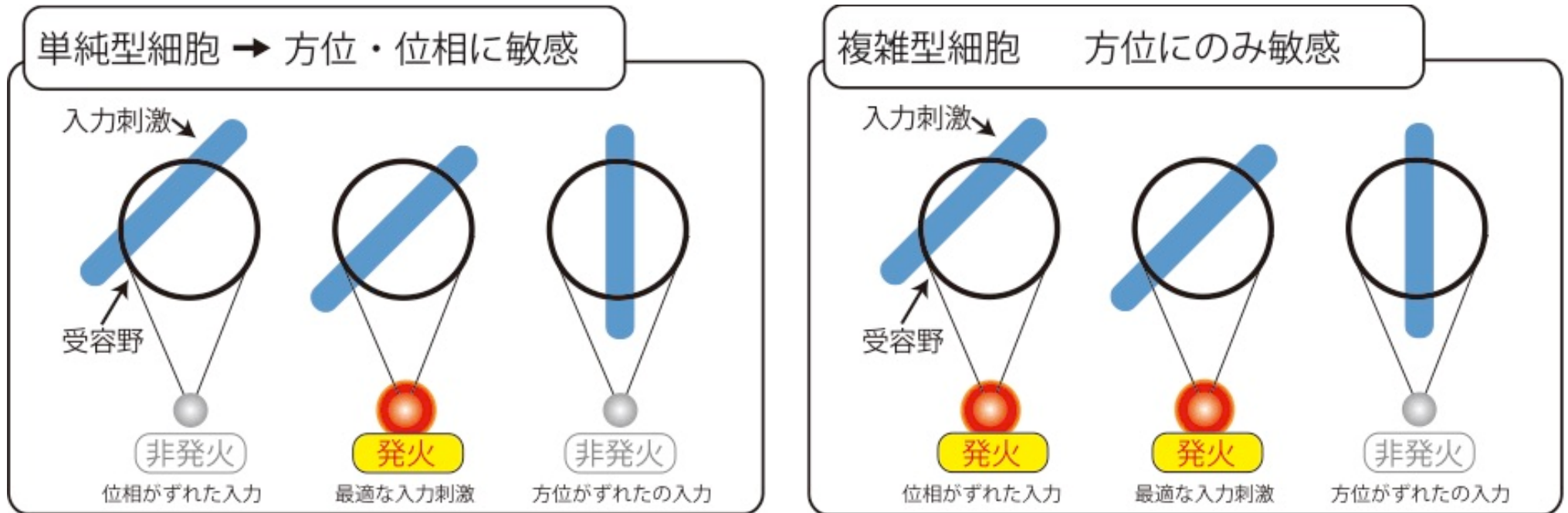
初期視覚野 (V1野) の受容野

- 1個の細胞の観えている範囲 (受容野)は狭い
→ **情報処理範囲が制限**
- 空間フィルタ演算で記述可能
- 視覚野中の局所的な
“エッジ” や “線分” 成分に反応
- **単純型細胞**(simple cell)と
複雑型細胞(complex cell)が存在



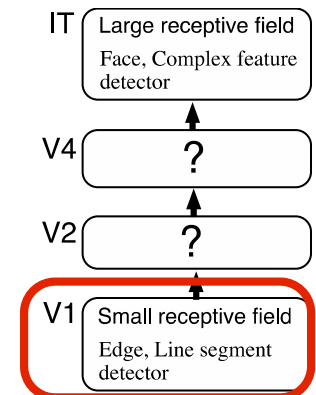
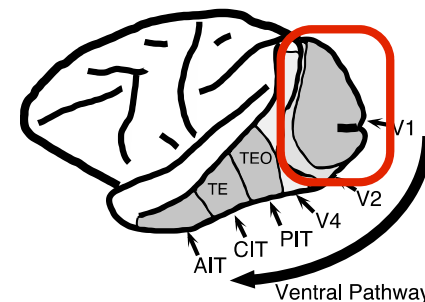
<http://ohzawa-lab.bpe.es.osaka-u.ac.jp/resources/text/KisokouKoukai2009/Ohzawa2009Koukai04.pdf>

初期視覚野の性質

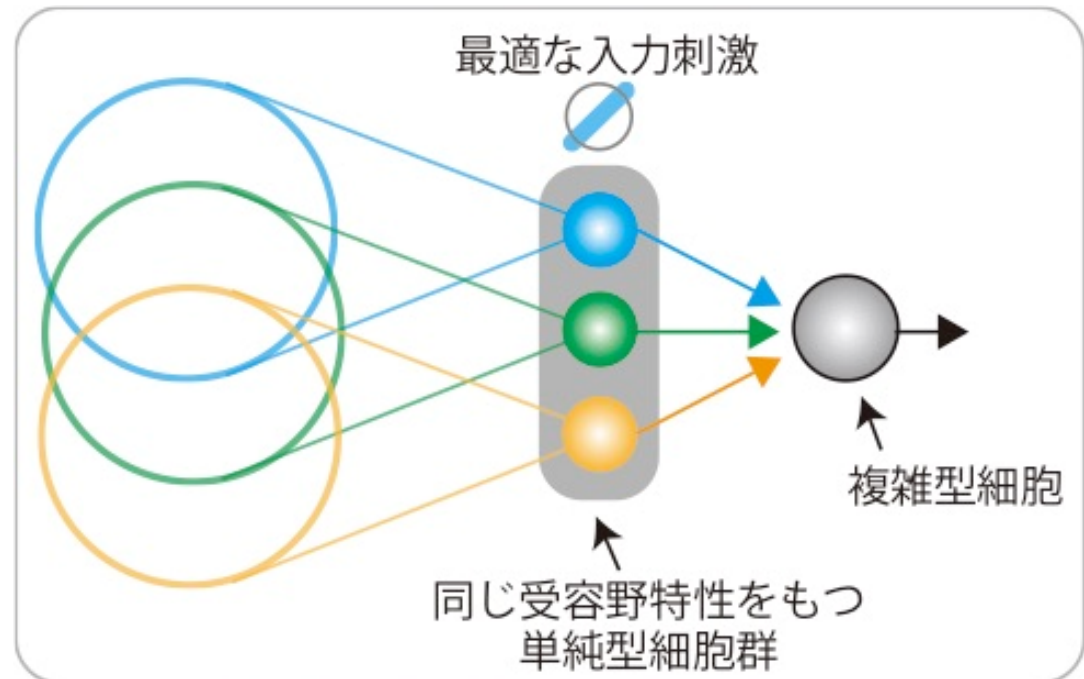
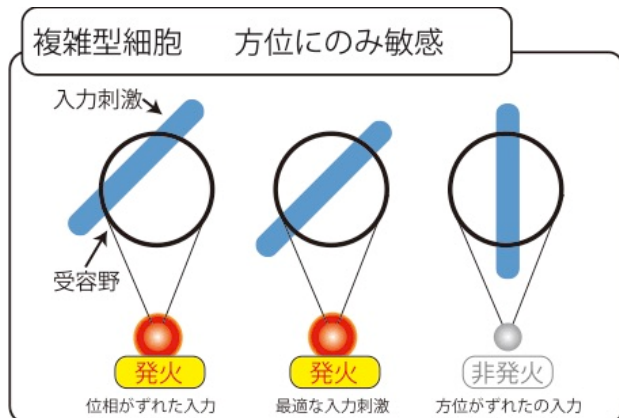
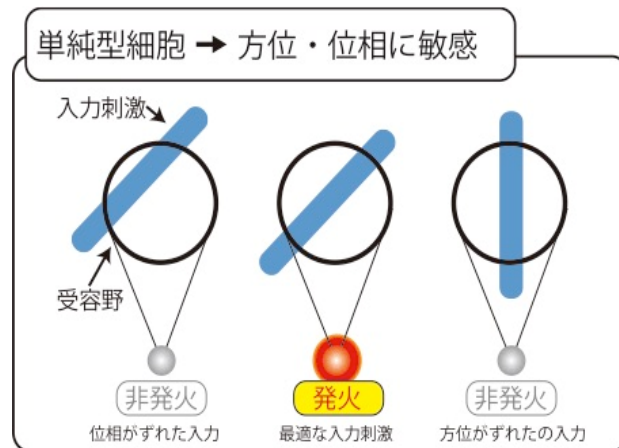


□ 線分やエッジなどの成分に反応

- Simple cell: 方位, 位相に敏感
- Complex cell: 位相には許容的

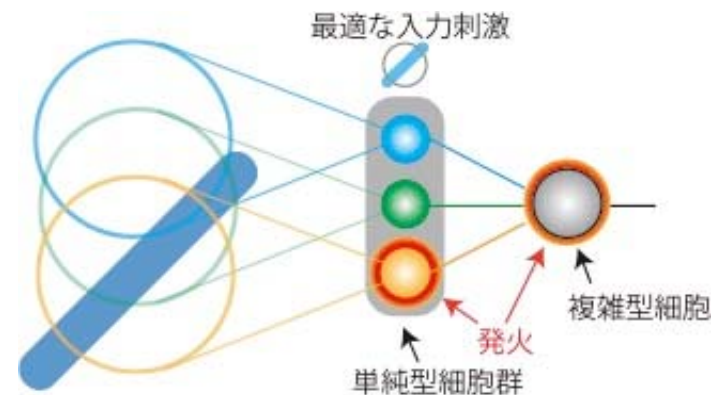


Hubel-Wiesel の階層仮説

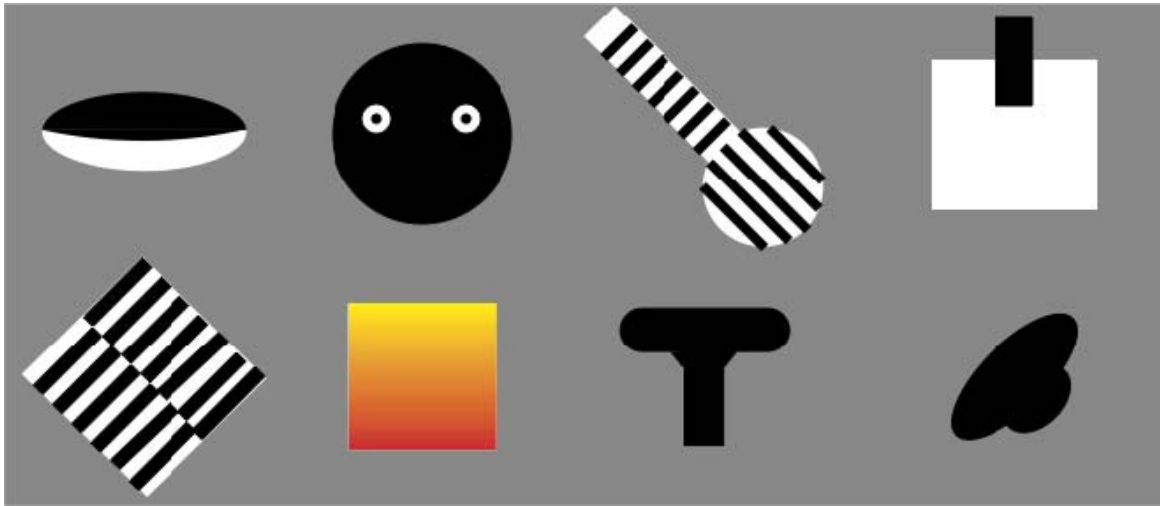


Hubel & Wiesel の階層仮説

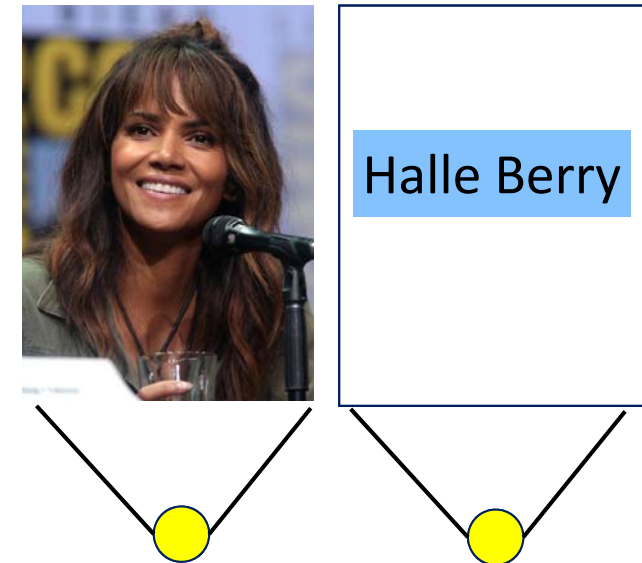
- 単純型細胞の出力合成で、
複雑型細胞出力が説明可能
[Hubel & Wiesel 59]



高次視覚野の性質

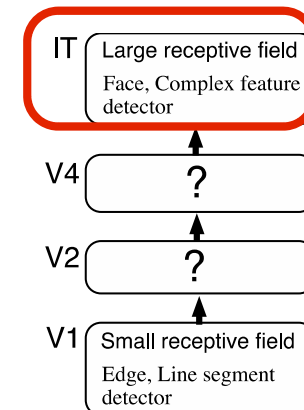
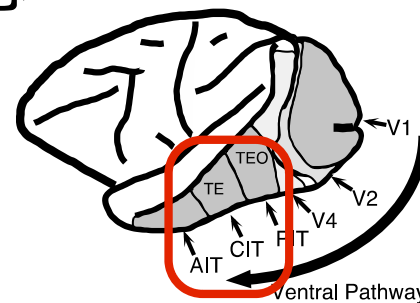


(Kobatake & Tanaka 94 を改変)



“Yes, It’s Halle Berry”

- 巨大な受容野
- 中程度に複雑な特徴に反応
- 顔細胞の存在
- 分散表現
- 時空間的な変化に許容的



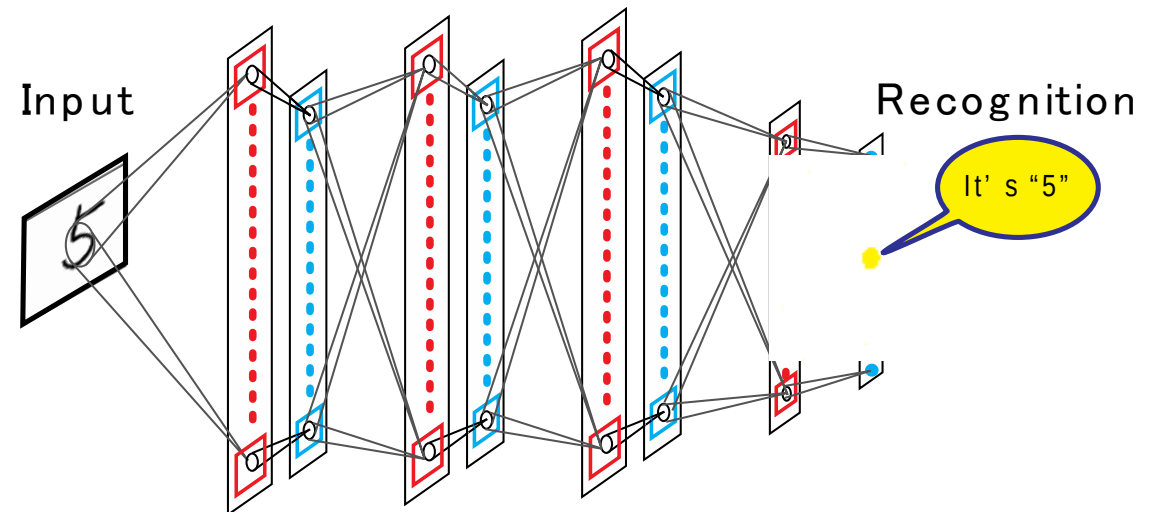
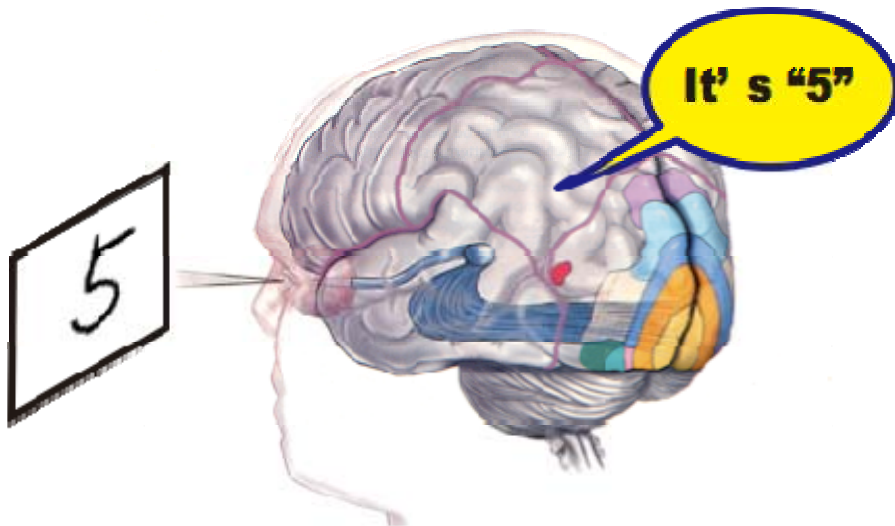
3 時間目

AIの視覚システムを作るには？

ディープラーニング

ディープラーニング(深層学習)とは？

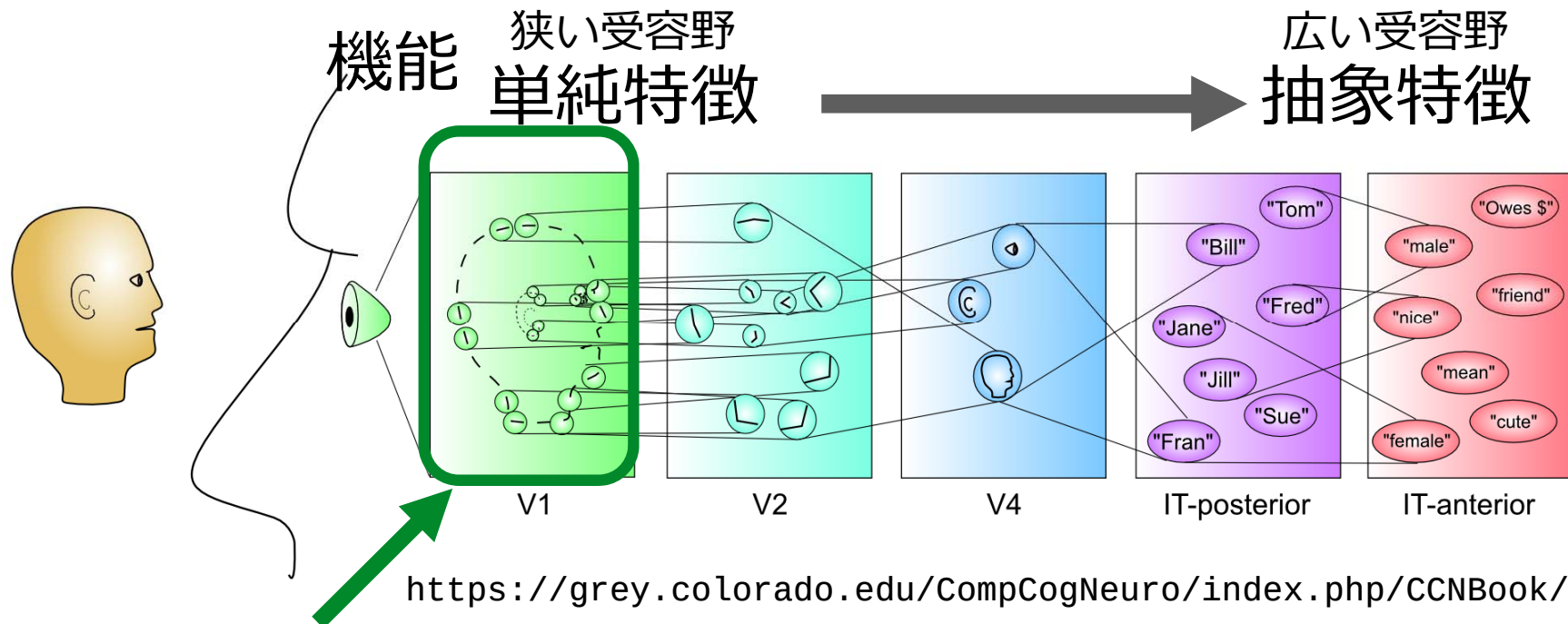
- ニューラルネットモデルを用いた機械学習(AI)技術
 - 脳の働きを模した構造と学習方式
 - 深い階層構造を持つことが特徴
- 第3次AIブームの中核技術
- 画像処理, 音声処理, テキスト処理などの分野で無類の強さを発揮中



Neocognitron/CNN [Fukushima 80]

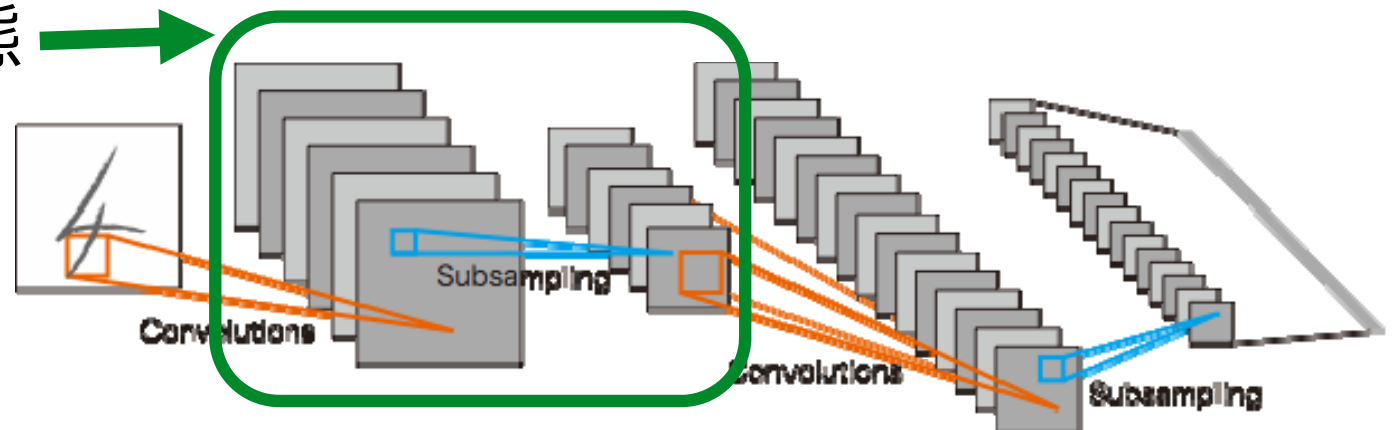
Modified from <http://thebrain.mcgill.ca/>

生理学的な知見とCNN



単純型/複雑型細胞

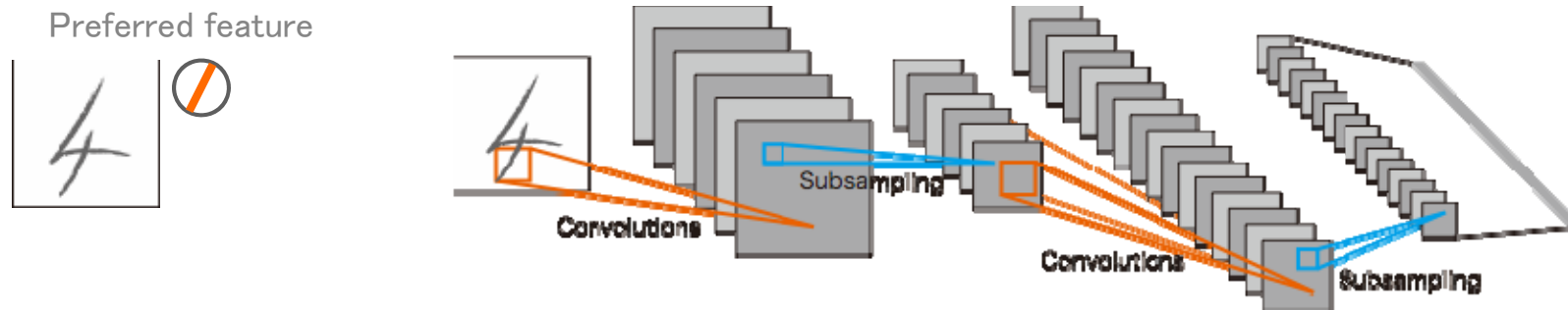
ハードウェアの実態



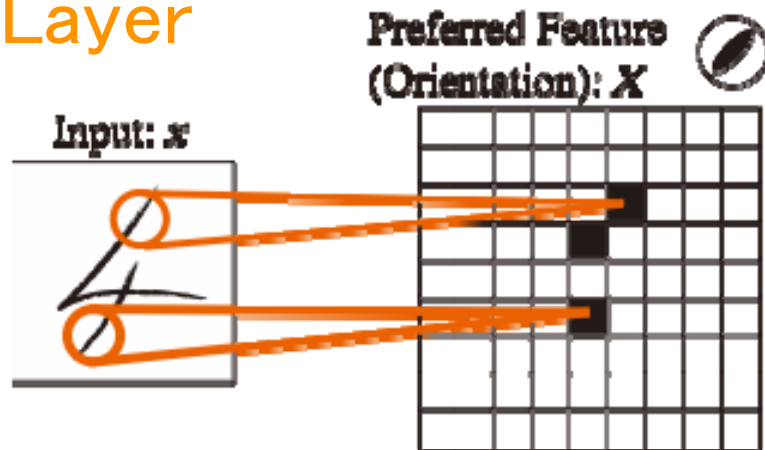
Convolution Neural Network: CNN, Neocognitron

CNNの動作原理

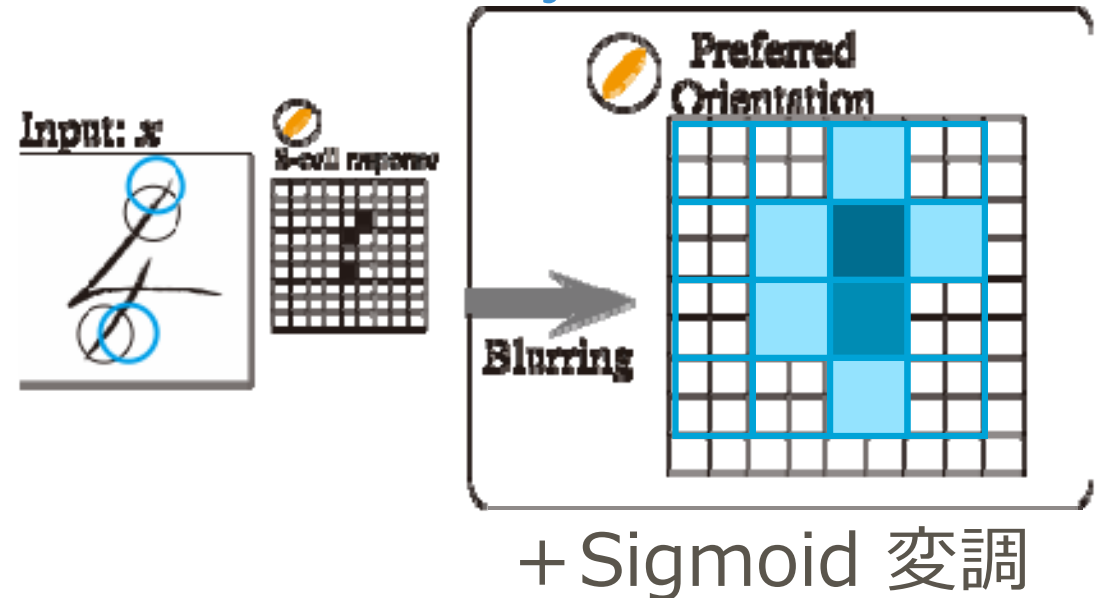
- 局所特徴抽出(畳み込み) + 変形に対する不変性(プーリング)



S-cell (Conv.)
Layer

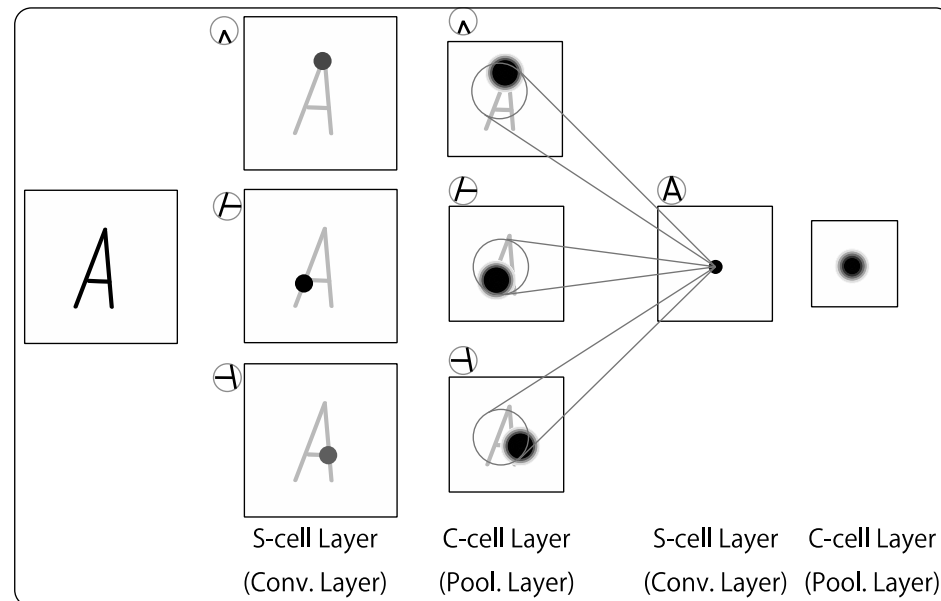
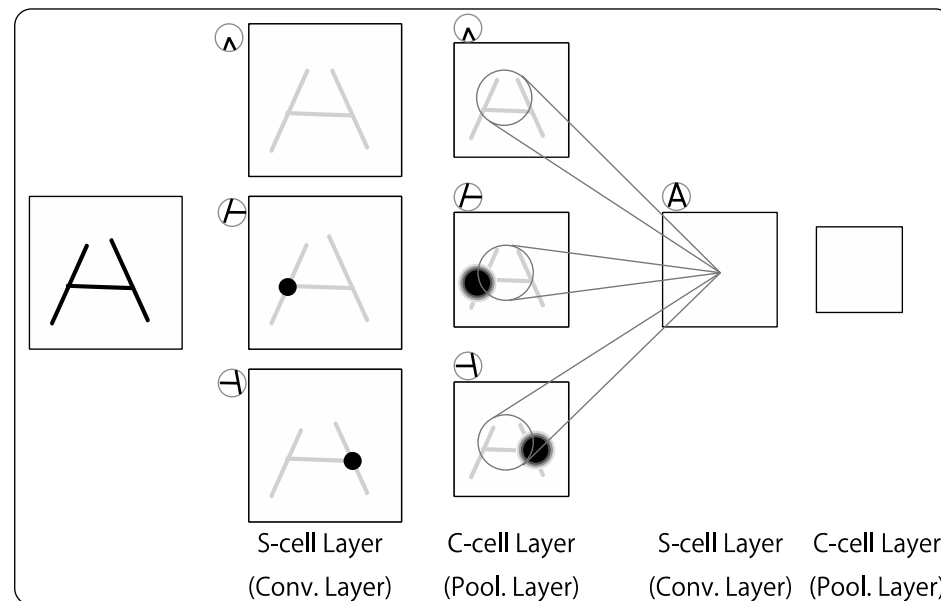
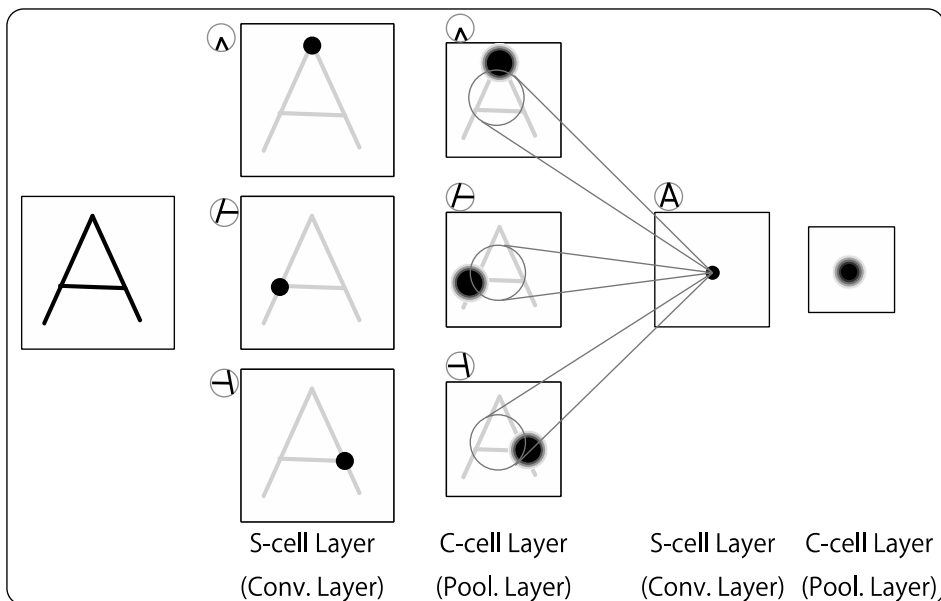


C-cell (Pool) Layer



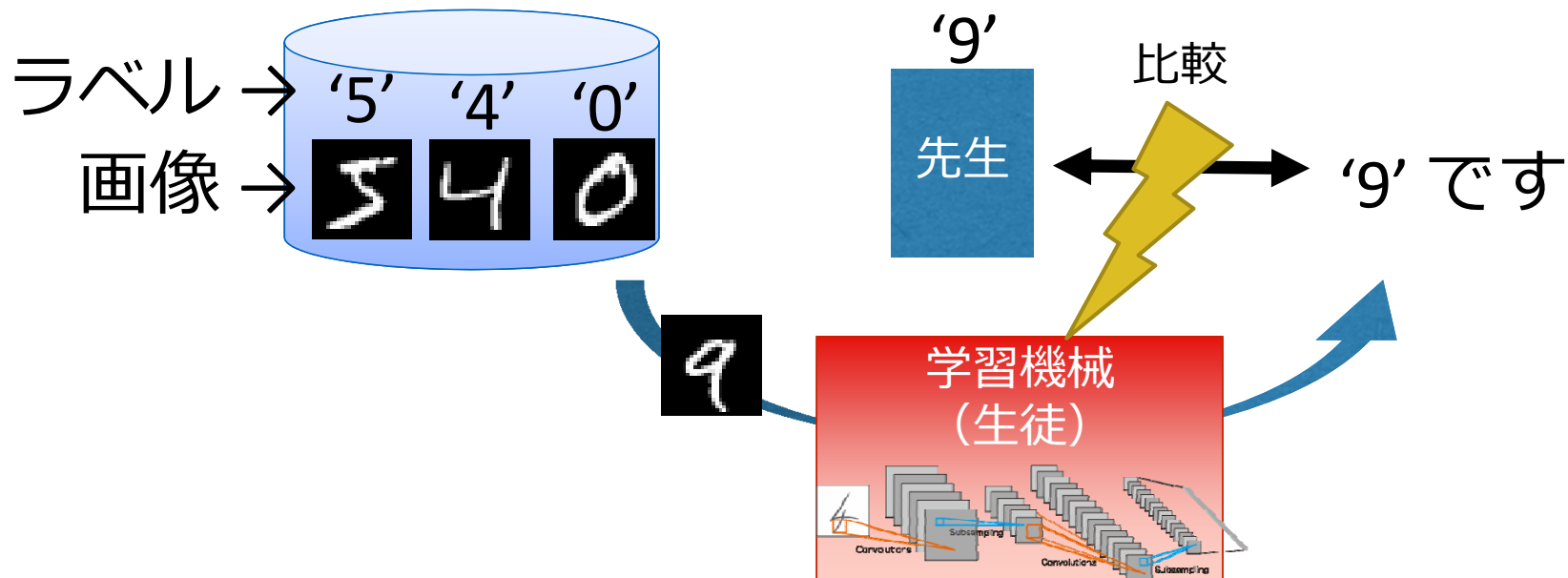
CNNの動作イメージ

局所特徴抽出(畳み込み) + 変形に対する不変性(プーリング)



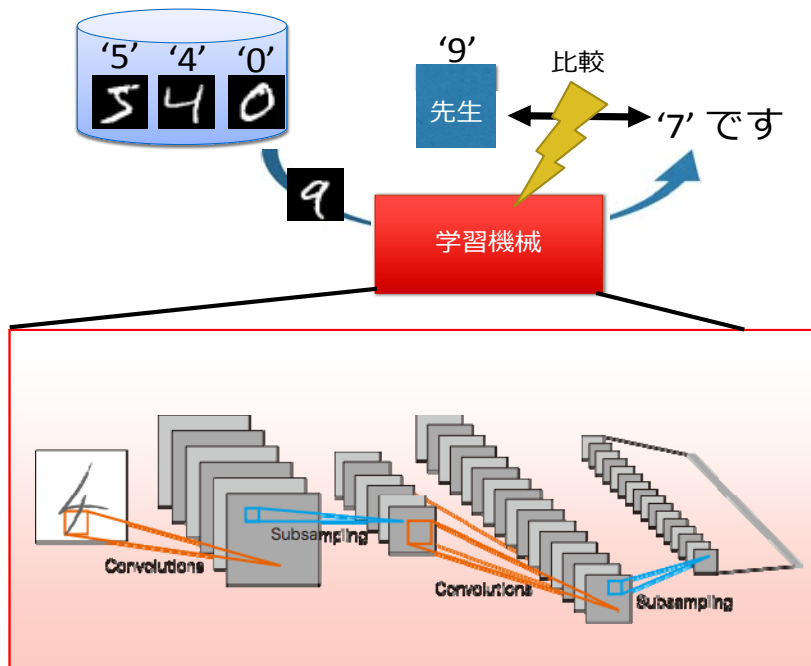
統計的機械学習の枠組み

- 識別問題を例にCNNの学習を考えてみる
 - 入出力をマッピングする関数がほしい
ただし関数の具体的な形はわからない.
 - ユーザー側が持っているのは(大量の)入出力データ
 - 正解データ（先生）との比較結果を学習機械内部の
パラメータに反映



CNNの学習とは

- “学習” は畳み込みフィルタの係数(w)の調整
- 入出力のマッピングが合致するように
“先生” と機械が同一の答えを出すように調整
- 先生の答えと生徒の答えの差異→ロス関数 $E(w)$



ロス関数の例

$$E(w) = \frac{1}{2} \sum_n \|t_n - y(x_n; w)\|^2$$

先生の答え

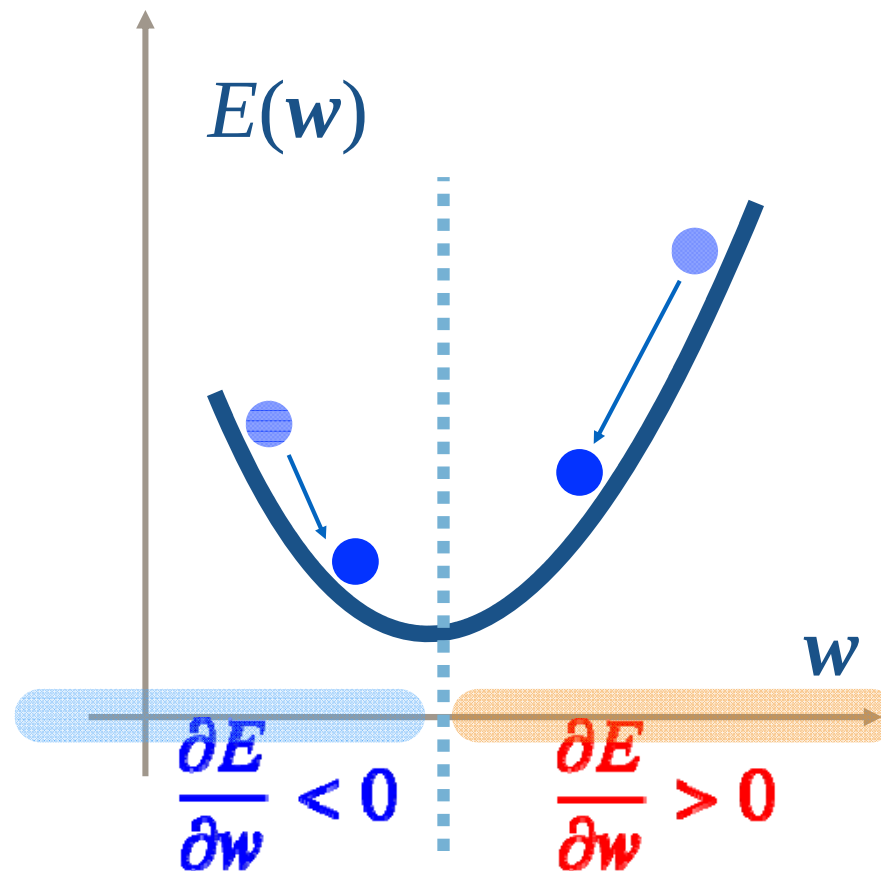
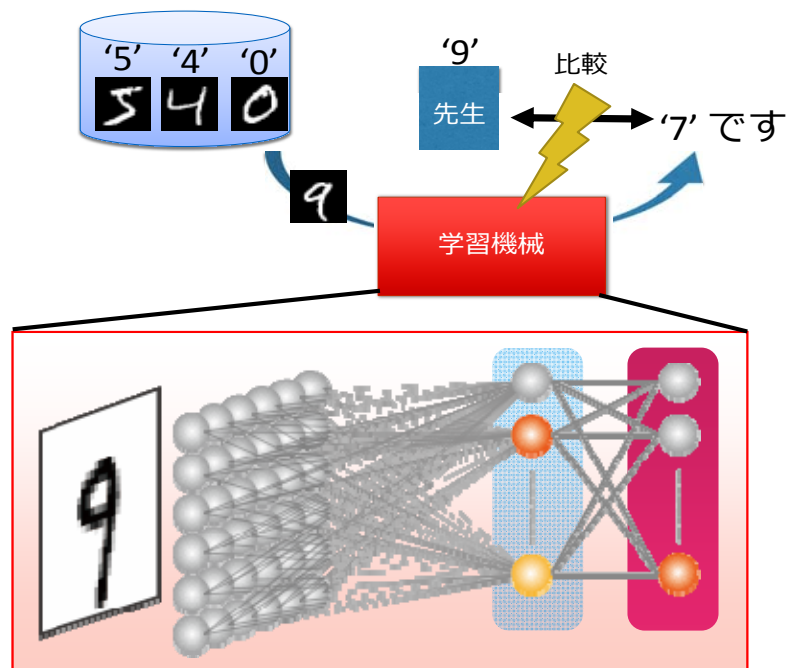
機械の答え

$E(w)$ 小さくする w を求めたい

勾配法を用いた学習

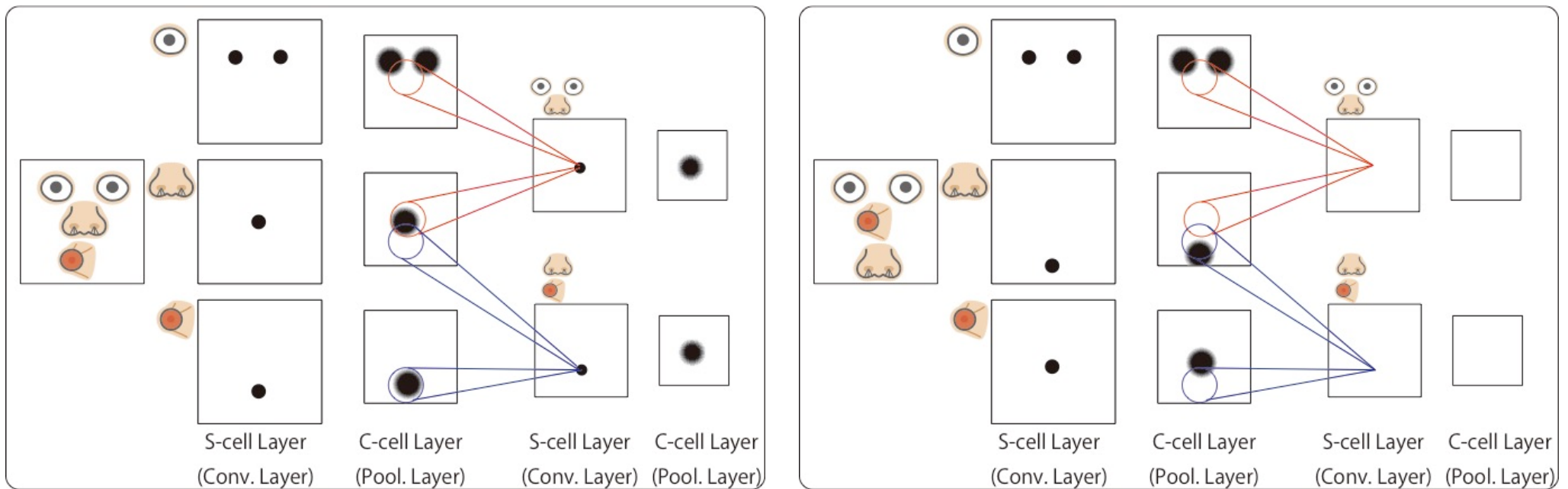
- “学習” はパラメータ w の調整
- ロス関数 $E(w)$ を最小化するために “勾配” を使う

勾配法のパラメータ更新則 $w^{t+1} = w^t - \eta \frac{\partial E(w)}{\partial w}$



CNN の動作イメージ

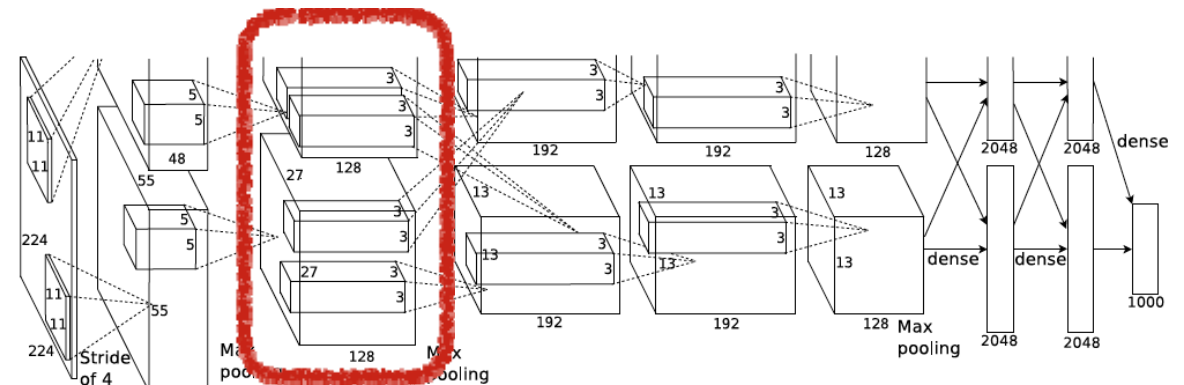
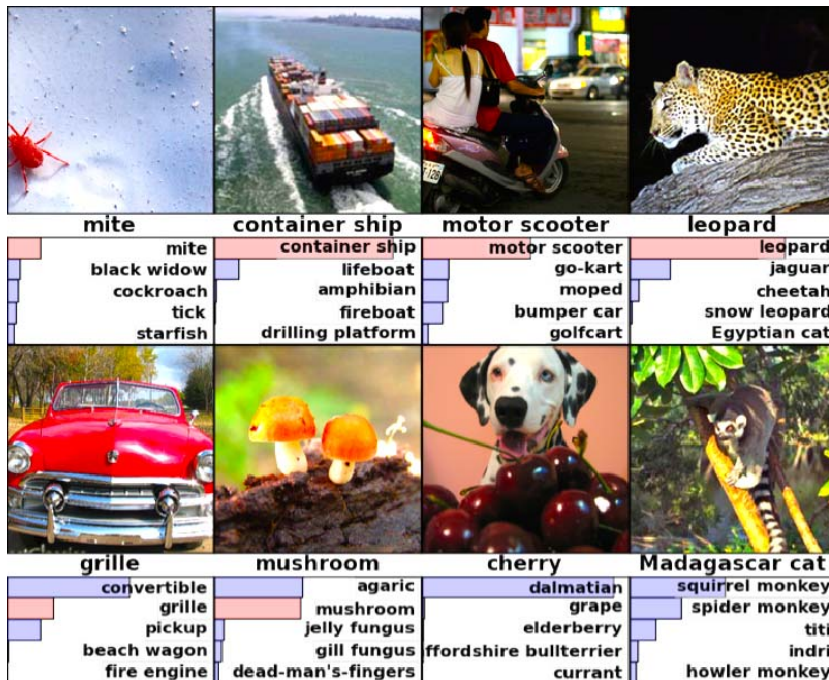
- CNN による顔認識タスク
“福笑い” を例に考えてみる



特徴の**組み合わせの位置関係を保持**しながら
新たな特徴を抽出していく

CNN 応用事例: 一般物体認識

- AlexNet: ILSVRC 2012 の勝者
 - ビルディングブロック: Conv + Pool が基本
- 1000 カテゴリ × 約1000枚の訓練画像



Krizhevsky et al. NIPS 2012

SIFT + FVs: 0.26 test err.

DCNN: 0.16 test err.

まとめ

まとめ

□ 1 時間目

古生物学的な視覚システムである側抑制について学び、Mach-band 現象といった錯視が説明できそうなことがわかった.

□ 2 時間目

フィルタ処理を用いた画像処理の概要と、視覚システムとの関わりについて学んだ.

□ 3 時間目

視覚システムを模倣した畳み込みニューラルネット (CNN) を構築し、数字文字認識を行った.