

自己組織化神経回路モデル (4)

- 論文紹介:

B.A. Olshausen, D.J. Field

Emergence of simple-cell receptive field properties by learning a sparse code for natural images, *Nature*, vol. 381, pp.607-609, 1996.

- 研究の動機:

1. 眼や脳が視覚情報をどう処理しているか .
2. デジタル画像の効率のよい表現方法 (情報圧縮)

- キーワード

- 過完備基底によるスパース表現, 自然画像, 独立成分分析, V1 の単純型細胞
- 各画像を, 入力信号空間の構造を捕えた基底の線型重み和で表現 .

- 例:

$$I(x, y) = \sum_{i=1}^{192} a_i \phi_i(x, y), \quad x = 1, \dots, 16, \quad y = 1, \dots, 16 \quad (1)$$

- 基準 1: 再現性

もとの各画像 $I(x, y)$ を再構成できる基底 $\phi_i(x, y)$, $i = 1, \dots$ を求めたい .
このようなモデルを生成モデルと呼ぶことがある .

- 基準 2: スパース性

再構成の際, できるだけ, 使う基底の数を少なくしたい .

$a_i, i = 1, 2, \dots$ のほとんどは, 値 0 になってほしい .

- 基準 3: 独立性 (ここでは直接的には使わない)

各 $a_i, i = 1, 2, \dots$ 間の相関をなくしたい . 同時エントロピー $H(a_1, a_2, \dots, a_n)$ は保ったまま, 個々のエントロピーの和 $\sum_i H(a_i)$ を小さくしていく . このような a_1, a_2, \dots を見つける . 数学的性質: 例えば a_1, a_2 間に依存性があれば, a_1 の値がわかれば a_2 の値がある程度, 推測できる . これは $H(a_1, a_2) < \sum_i H(a_i)$ ということ . 確率変数 a_1, a_2, \dots, a_n 間に依存性があれば, $H(a_1, a_2, \dots, a_n) < \sum_i H(a_i)$ となる . これを a minimum-entropy code と H.B. Barlow は名付けた .

- 目的関数

$$E = -[\text{preserve information}] - \lambda [\text{sparseness of } a_i] \quad (2)$$

$$\frac{1}{2} \sum_{x,y} \left[I(x, y) - \sum_{i=1}^{64} a_i \phi_i(x, y) \right]^2 + \lambda \sum_i S\left(\frac{a_i}{\sigma}\right) \quad (3)$$

これを最小にするような $a_i, \phi_i(x, y)$, $i = 1, 2, \dots$ を求めたい . 学習が終わった段階では, 個々の入力画像に対し, a_i は異なった値をとるが, どの入力画像に対しても固定した基底 $\phi_i, i = 1, 2, \dots$ を使って表現する .

- ここで $S(x)$ というのは, スパース性を促進する, 例えば以下のようなもの .

- * $S(x) = -e^{-x^2}$

- * $S(x) = \log(1 + x^2)$

- * $S(x) = |x|$

- 2つのダイナミックス .
 - * 速い a_i のダイナミックス . ニューロン活動に対応 .
 - * 遅い ϕ_i のダイナミックス . 学習 (結合係数の変化) に対応 .
- 学習 (自己組織化)

目的関数 E を a_i で微分してみる .

$$\begin{aligned}
 -E &= -\frac{1}{2} \sum_{x,y} \left[I(x,y)^2 - 2I(x,y) \sum_{i=1}^n a_i \phi_i(x,y) + (a_1 \phi_1 + \dots + a_n \phi_n)^2 \right] - \lambda \sum_i S\left(\frac{a_i}{\sigma}\right) \\
 &= \frac{1}{2} \left[-\sum_{x,y} I(x,y)^2 + 2 \sum_{x,y} I(x,y) \sum_{i=1}^n a_i \phi_i(x,y) - \sum_{x,y} (a_1 \phi_1 + \dots + a_n \phi_n)^2 \right] - \lambda \sum_i S\left(\frac{a_i}{\sigma}\right) \\
 -\frac{\partial E}{\partial a_i} &= \sum_{x,y} I(x,y) \phi_i(x,y) - \sum_{x,y} \sum_j (a_j \phi_j \phi_i) - \frac{\lambda}{\sigma} S'\left(\frac{a_i}{\sigma}\right)
 \end{aligned}$$

したがって, E を減らす方向に a_i を動かすには $\dot{a} = -\frac{\partial E}{\partial a_i}$ とすればよい .

- 図 1: 自然画像からとってきた大量の 8×8 データに対して主成分分析した結果 . 64 個すべての基底 (固有ベクトル) を表示 .
- 図 2: コスト関数の模式図 . $S(x) = \log(1 + x^2)$. 各 a_i はなるべく 0 になろうとする .
- 図 3: この論文で提案しているアルゴリズムが確かに動くことを, 人工的に作った画像で確認 . あらかじめ, ある基底集合の重みづけで画像を生成 . 3 種類の基底を試しており, それが a, b, c に対応 . もとの基底表現が自己組織的に獲得できたことに注目 .
 - a) 各ピクセルごと独立にランダムな値を割り当て (大きな値は減多にとらないようモデル化) .
 - b) フーリエ変換した領域で a) のようなものを作成 .
 - c) a) b) では互いの基底は直交していたが, c) では, 互いに直交していない Gabor 関数 (基底) の重ねあわせで画像をランダムに生成 .
- 図 4: a. 学習により獲得した基底 $\phi_i(x,y)$, $i = 1, \dots, 192$ を表示 .

入力: 16×16 -pixel の画像パッチ . もとの画像は 512×512 の自然画像 10 枚 . 様々な傾き, 位置, 空間周波数の基底が得られたことに注目 .

 - b. 各場所のスポット入力 (1 ピクセルだけ on?) に対する再構成時の a_i の値を表示 . $16 \times 16 = 256$, それぞれの場所で試す .
 - c. どんな空間周波数の基底が得られたか, その分布 .
 - d. 自然画像のパッチを表現したときの a_i の確率分布 (ヒストグラム) .

実線: 学習済の基底を使った場合 . $H(a_i) = 4.0$

点線: ランダムに生成した基底 . $H(a_i) = 4.6$

確かにエントロピーが下がっている ($\sum_i H(a_i)$ をできるだけ小さくするのが目的だった) .

P.Földiák (1990) の論文との関連：基本的に同一の思想をもった研究

- 同じ点：

各々の入力に対し，自己組織的に（教師なし学習により）内部表現を形成する．

2つのダイナミックス：

1. 各入力を変現する（ニューロン活動のダイナミックス）速いダイナミックス．
2. 変現するための基底（結合係数）を学習により獲得する遅いダイナミックス．

- 若干異なる点：

モデル

Földiák (1990) は神経回路モデル．

Olshausen & Field (1996) は神経回路モデルではないが，神経回路的な解釈はできる．

再構成（生成）性

Földiák (1990) では再構成は考えていない．再構成してみると，どういう画像が得られるかは，示していない（見てみたい）．

Olshausen & Field (1996) は，各入力を基底の線型重み和で再構成．

入力画像

Földiák (1990) では文字が描かれた画像の 0,1 パターン．

Olshausen & Field (1996) では自然画像の濃淡値．視覚一次野のニューロンの反応選択性と比較したいこともあり，画像全体ではなく一部分（ 16×16 のパッチなど）．

- 異なる点：

線型・非線型モデル

Földiák (1990) は非線型．

Olshausen & Field (1996) は各基底の線型の重み和で画像を表現．

基底を使った表現の仕方

Földiák (1990) の場合，入力を表現する際，各基底に対するの重みづけはほぼ 0,1（ニューロンの活動）．

Olshausen & Field (1996) では a_i はスパースとは言ってもアナログ値（このアナログ値を実際の神経活動でどう表現しているのかは，よくわからない）

エントロピーを使った評価方法

Földiák (1990) では，入力に与える各アルファベット・記号の出現頻度のある英文と同じにして，エントロピーによる評価をおこなった．符合化の前後で情報量はほぼ落ちていないが，各ビットのエントロピー（ a_i に対応）の和は低く押さえていること，各 a_i が独立に近くなっていることを示した．

Olshausen & Field (1996) は入力が自然画像の一部分で，しかも濃淡値であり，可能な画像の数は膨大なものになることから，Földiák (1990) のような解析はおこなえない．そこで，どの a_i も，その確率分布が同じ形をするはずであると考え，すべての $a_i, i = 1, 2, \dots$ を同一視して，その確率分布のエントロピーを計算した．その結果，エントロピーが確かに小さくなっていることを示した．