1F3-2

スパース性に基づく受容野の変数分離指標の推定

Sparse Estimation of Receptive Field's Separability Index

四倉 晋平 *1　大森 敏明 *2　五十嵐 康彦 *1　岡田 真人 *1
Shimpei Yotsukura　Toshiaki Omori　Yasuhiko Igarashi　Masato Okada

*1東京大学大学院新領域創成科学研究科
Graduate School of Frontier Sciences, The University of Tokyo

*2神戸大学大学院工学研究科
Graduate School of Engineering, Kobe University

Neuron encodes spatiotemporal input into spikes train with its receptive field. We can distinguish two types of neurons, examining whether each receptive field is spatiotemporally separable or not. In this study, we propose a method to estimate the Separability Index (SI), the degree of separability, by means of sparse linear regression with basis functions. Using numerically simulated data, we show that the proposed method achieves more accurate estimation of the SI than simple trial average used in conventional method.

1. 研究背景と概要

神経細胞は入力された時空間情報を，スパイクの発火として，出力に変換することにより，情報処理を行っている．この入出力関係は，受容野とよばれる時空間フィルタによって強く特徴づけられている．ここで受容野が，時空間的に変数分離化可能か否かを調べることによって，神経細胞の機能を弁別することが可能である．

この受容野の変数分離可能性を表す指標として，変数分離指標 (Separability Index, SI) が知られている．しかし，SI を求める際には，データの不足に起因するノイズが重畳され，SI から正しく変数分離可能か否かを決定することができない．

本研究では，ノイズが重畳された受容野の実験データから，スパース推定に基づき，受容野の変数分離指標を精度良く推定する方法を提案する．

2. 神経細胞の受容野とスパイクトリガー平均

神経細胞は，入力された情報を，受容野とよばれるフィルタをによって変換し，スパイク出力を放出することにより，情報処理を行っている．したがって，神経細胞の入出力特性は，受容野によって強く特徴づけられる．

本研究では，この受容野のモデルのうち，一次視覚野の単純型細胞のモデルとして，以下のように定義される時空間ガボールフィルタ $g(x,t)$ を扱う [Palmer 87, DeAngelis 95].

$$g(x,t) = \exp\left\{-\frac{(a^2+\gamma b^2)}{\sigma^2}\right\}\cos\left(2\pi\frac{a}{\lambda}\right) \tag{1}$$

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ t \end{pmatrix} \tag{2}$$

ただし，σ,γ，λ は定数である．図 1 にガボールフィルタの形状の例を示す．このガボールフィルタ $g(x,t)$ を用いて，時刻 t における単純型細胞の発火率，すなわち単位時間当たりのスパイクの出力数は

$$r(t) = \left[\int dx \int d\tau g(x,\tau) I(x,t-\tau)\right]_+ \tag{3}$$

と記述される．ただし，$[\cdot]_+$ は，線形閾値関数である．

連絡先: yotsukura@mns.k.u-tokyo.ac.jp

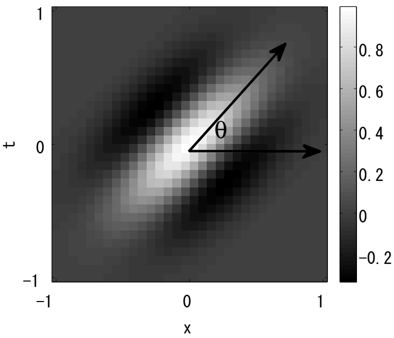

図 1: ガボールフィルタの形状の例．$\theta = 0°$ の場合には，変数分離可能であり，$\theta \neq 0°$ の場合には，変数分離不可能である．この図では，$\theta = 45°$ となっており，変数分離不可能な場合に対応する．

この受容野を実験的に求める方法として，スパイクトリガー平均 (Spike-Triggered Average, STA) が知られており，生理実験に基づいた神経細胞の受容野推定に用いられている [Simoncelli 04, Paninski 04]．まず，対象とする神経細胞へ入力として，$I(x_i,t_k)$ をガウスノイズに従うように与える．このような状況で STA を，

$$g_{STA}(x_i,t_j) = \frac{1}{K}\sum_{k=1}^{K} r(t_k) I(x_i, t_k - t_j) \tag{4}$$

と定義する．上記のように STA を求めることで，$K \to \infty$ のとき，受容野に一致することが知られているが [Paninski 04]，実際の実験では実験回数 K が有限のため，STA にはノイズが重畳される．

3. 受容野の変数分離指標

受容野が，時空間的に変数分離可能であるか否か，すなわち $g(x,t) = g_x(x)g_t(t)$ と記述できるかを調べることにより，その神経細胞の機能を弁別することができる [DeAngelis 95]．例えば，図 1 に示すように，ガボールフィルタにおいて，$\theta = 0$ の場合には，時間的に移動しない，ある角度の線分を認識する，

方位選択性の細胞と決定できる．対して，$\theta \neq 0$ の場合には，運動する線分を検出する，方向選択性の細胞に分類される．

この変数分離可能性を示す指標として，以下のように求められる変数分離指標 (Separability Index, SI) が用いられている [Depireux 01]．まず受容野を行列化し，特異値分解を行う．すなわち $G_{ij} = g(x_i, t_j)$ なる行列 G に対して，$G = \sum_i s_i \boldsymbol{X}_i \boldsymbol{T}_i^t$ となるように，変数分離可能な行列の和で書き下す．ただし，$\boldsymbol{X}_i$，$\boldsymbol{T}_i$ は，i 番目の特異ベクトルである．

このような状況で SI は

$$\text{Separability Index} = \frac{s_1^2}{\sum_i s_i^2} \tag{5}$$

と定義される．ただし，s_1 は s_i のうち，最も絶対値の大きいものとする．したがって SI の値は，受容野のうち最も寄与の大きい変数分離可能な行列によって，どの程度受容野を表現できるかの値である．もし受容野が変数分離可能であるならば，SI= 1 となる．しかし，受容野の情報として，STA の計測データを用いるため，ノイズの影響を受け，図 2 に示すように正しく SI を求めることができない．

4. 提案法

本研究では，ノイズが重畳された STA データから，元の受容野の SI を精度良く求める手法を提案する．まず生理実験により STA のデータ点が得られたとする．$g_{STA}(x,t)$ を基底関数 $\{f_m\}_{m=1\cdots M}$ と実数 $\{a_m\}_{m=1\cdots M}$ を用い，

$$\hat{g}(x,t) = \sum_{m=1}^{M} a_m f_m(x,t) \tag{6}$$

と．線形回帰の枠組みで推定を行う．すなわち係数 $\{a_m\}$ を STA のデータ点に基づいて決定し，STA を推定する．

式 (6) の回帰では，どのような基底関数を用いるかが回帰の精度を決定する．視覚野の神経細胞の受容野は，特定の周波数刺激に強く反応することが知られている．したがって，2 次元のフーリエ基底 $\cos(\pi ux + \pi vt), \sin(\pi ux + \pi vt)$ を用いて，周波数特性を表現する．

フーリエ基底を用いるうえでどの次数までの基底を用いて回帰するのが良いのかは明らかではない．線形回帰の際に冗長な基底関数を用いると，ノイズまで回帰モデルに組み込むため，オーバーフィットが生じ，推定の精度が悪化することが知られている．したがって測定ノイズによる影響を受けない形で推定するため，STA に対して本質的な基底関数のみを抽出して回帰を行う必要がある．このような基底関数の抽出を行う方法として，L1 正則化がある [Tibshirani 96]．L1 正則化と STA のデータ点を用い，次の目的関数

$$E = \sum_{i=1}^{N}\sum_{j=1}^{N} \{g_{STA}(x_i,t_j) - \hat{g}(x_i,t_j)\}^2 + \sum_{m=1}^{M} \lambda_m |a_m| \tag{7}$$

を最小化するような係数を選択する．目的関数 E の第一項は回帰モデルと STA の測定データの精度を測る項であり，第二項は係数の絶対値が大きくなることを防ぐ重みづけ罰則項である．L1 正則化ではデータの回帰に適さない冗長な基底関数は，対応する係数が 0 になる傾向が強い．そのため L1 正則化により，回帰の誤差は減少させると同時に，本質的な基底関数を抽出し適切な回帰モデルを選択するという，スパース推定の枠組みを実現できる．

L1 正則化を適用するうえで，係数の絶対値に対する罰則の重み λ_j を決定する必要がある．STA のデータにはノイズが重畳されるため，フーリエ基底の高周波成分がノイズの影響により過剰に選択される可能性があると考えられる．そこで高周波のフーリエ基底が抽出されることに強い罰則をかける．周波数が d のフーリエ基底に対応する重みは周波数に比例させ $\lambda_j = d\lambda$ とする．ただし λ は正の定数で，10 分割交差検定を用いて，データから自動的に決定する．

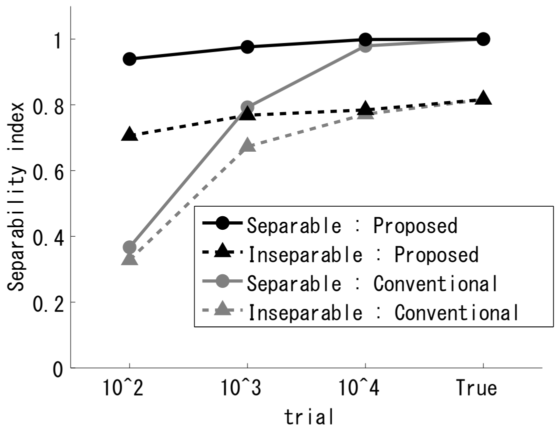

図 2: 提案法による SI の推定結果．横軸は実験回数を表し，右端は真値である．灰線は従来法で求めた SI であり，黒線は提案法により求めた SI である．丸は変数分離可能な $\theta = 0°$ の場合，三角は変数分離不可能な $\theta = 15°$ の場合の結果である．提案法により，SI の真値を精度よく求めることができた．

5. 実験結果

図 2 に，単純型細胞の数値実験データに，提案手法を適用した場合の，推定された SI の結果を示す．図の横軸は実験回数を表し，右端は真値である．灰線は従来法で求めた SI であり，黒線は提案法により求めた SI である．丸はガボールフィルタの $\theta = 0°$ の変数分離可能な場合，三角は $\theta = 15°$ の変数分離不可能な場合の結果である．

従来法は実験回数が少ない場合に，ノイズの影響により，真値と大きく異なる値になっており，従来法により求めた SI では，変数分離可能な細胞と不可能な細胞を分別することができない．対して提案法を適用した結果，SI の真値を精度よく求めることができていることが分かる．提案法により，神経細胞の実験回数が少ない場合にも，神経細胞の機能を弁別できることが示された．

参考文献

[DeAngelis 95] G. C. DeAngelis, I. Ohzawa, and R. D. Freeman, Trends Neurosci., 1995.

[Depireux 01] D. A. Depireux, J. Z. Simon, D. J. Klein and S. A. Shamma, J. Neurophysiol., 2001.

[Palmer 87] J. P. Jones and L. A. Palmer, J. Neurophysiol., 1987.

[Paninski 04] L. Paninski, Network, 2004.

[Simoncelli 04] E. P. Simoncelli, J. Pillow, L. Paninski and O. Schwartz, Gazzaniga (ed.), MIT Press, 2004.

[Tibshirani 96] R. Tibshirani, J. Royal. Statist. Soc. B, 1996.