

隠れ状態とマルコフランダム場 講義録

講 師：岡田 真人

レポーター：坂谷智也・土肥 英三郎

1999年8月24日

この講義で最低限理解してほしいこと

- これから幾つかややこしいことを説明するわけですが、何故そのようなことを考えるのかという動機を理解して下さい。
- 結合 MRF モデル、隠れ状態、マルコフランダム場とは何か。
- 結合 MRF モデルの神経回路網へのインプリメンテーション。これは偏微分を知ってさえいればできます。

1 動機

1. 先程の講義で出てきた Lamme らの文脈的修飾の話を定量的に扱いたいということがあります。
2. 昨日までの話では重要な点が抜けています。昨日までの話というのは、画像が何らかの基底の線形和で表せるというものです。しかし、画像の生成プロセスは一般に非線形です。理由は色々ありますが、一番大きな理由は 3 次元の情報が網膜で 2 次元に落ちるという点にあります。例えば緑の面を赤の面が遮蔽している場合、それぞれの面は線形和で表せても、境界の部分では性質が急激に変化していて、線形和では表せません。非線形です。本当はこういうものを相手にしなければいけない。そのためにはエッジを取り出したり、それをもとに領域分割を行う必要があります。

Q: それは V1 でやるべきことですか？

A: 画像生成プロセスは一般には非線形なんですが、V1 だと受容野が小さいので、近似的に線形とみなすことができます。樺島先生の演習問題でも、画像を小さめに切って、それを見せて受容野を形成させている。つまり、緑と赤の面が同時に出てくるような状況をなるべく避けている。しかし我々の知覚は画像全体であり、それは非線形なプロセスによって生成されたものです。本当はこれを取り扱わないといけない。昨

日までの話では、まだこの問題が残っているわけです。これを実験的に扱おうとすれば、Lammeらの話になります。我々のグループも同じようなことをやっています。

3. 画像の話とは別に、情報のキャリアは何か、ということがあります。例えば同期が脳の中に物理的にあるかといえば、当然あるだろう。しかしそれが脳の情報処理として存在するのかどうかは分からない。情報のキャリアかどうかを確かめるためには、脳が行っている計算問題を考えて、その枠組みの中でスパイクの同期が必要かどうかを議論しなければいけない。

(Lammeの話に戻る)

V1には視覚の最終段階であるITとかTEへの順行性結合がある。そこから逆方向に逆行性結合もある。Lammeらは、彼らの示したV1の文脈的修飾はITやTEからの逆行性結合で起こっているだろうと言っています。しかし、これは良く考えると問題を含んでいる。

古典的受容野に対する刺激が同一でも、それが図にあるか、地にあるかで発火率が変わってくる(テキスト図1)。30-40ミリ秒遅れて図にある方が発火率が大きくなる。このメカニズムとして、Lammeらは側頭皮質からのフィードバックによるものではないかと言っている。この実験的な証拠も少しあります。

これを思考実験をしてみます(テキスト図2左)。V1のように受容野が小さく空間解像度が高い領野からITのように受容野が大きく空間解像度が低い領野に順行性結合が、逆方向には逆行性結合がある。この逆行性結合によって文脈的修飾が起きている、と彼らは言っています。

ここで問題なのは、解像度の低い情報を使って修飾を行うと、せっかくV1で保持していた解像度の高い情報が失われてしまうことになるのではないかと、ということです。

しかし、実際に文脈的修飾によって領域の境界部分に存在する解像度の高い情報が乱されているかといえば、そうはなっていない。

Q: 逆行性の影響が解像度の高い細胞の符号を変えるほどには強くないと考えてはいいかないのか？

A: そういう考え方もあると思います。実際にはポピュレーションコーディングのようにチューニングが緩やかで符号が0-1で決まっているわけではないので、この話を一般化するのには難しいと思います。

文脈的修飾がどれくらいの解像度かといえば、V1の解像度とほぼ同じです(テキスト図3)。つまりITの受容野よりずっと精緻な制御が行われている。場所をよく知らないものから信号を受けているのに、その影響は精緻になっている、という実験的な事実があるわけです。

これがどのようなことができるのかという問題をこれから議論していきます。結論から言うと、逆行性結合を適応的に切るような制御を考えるということです。これを結合MRFモデルで行う。

2 MRF モデル

今までの話は上から下の結合を切るということでしたが、それはまだやっていません。ここでは表面再構成という問題を考えます。まず先程の Lamme の話のように、明るさの異なる図と地を考えます。簡単のために 1 次元で考えるとハット状になります（白版で説明）。ここで、この 1 次元上の明るさを検出するセンサーがノイジーなので、得られるデータにノイズが加算されているとします。このデータに基づいて元の値を再構成するという問題を考える。

p1.m というプログラムを立ち上げて下さい。青い線がデータで、赤い線が推定値です。N を押しつづけると、赤い線がだんだん「なまってくる」ことが分かります。これが何をやっているかということこれから説明します。

まず、推定値 f_i はデータ d_i から余り離れてない方がいいので、 $\sum_i (f_i - d_i)^2$ の最小化を考える。ただ、これだけだと推定値とデータが同一のときに最小なので、推定といっても何もやらないことになる。何らかの加工を行いたい。

そこでもう 1 つ、画像は隣同士で近い値をとりやすい、ということを考える。これは隣接する推定値の差が小さくなるように、 $\sum_i (f_{i+1} - f_i)^2$ の最小化を考えればよい。

結局、この両方を小さくすることを考えることにします。つまり、ほどほどにデータを説明して、ほどほどに画像が本来持っているであろう性質を満たす、ということです。この 2 つを適当な比率で足し合わせたものを E と置きます。

$$E(f|d) = \frac{1}{2} \sum_i (f_i - d_i)^2 + \lambda \sum_i (f_{i+1} - f_i)^2$$

これを最急降下法で最小化するプログラムが先の p1.m です。 E は f_i の 2 次関数なので、下に凸の関数です。従って f_i を $-\partial E / \partial f_i$ に比例して更新してやれば E は最小値へ向かっていきます。

これを神経回路網でインプリメントすると次のようなアーキテクチャが現れます（白版の図）。定性的には、 f_i の作用（隣接に $+\lambda$ 、自分に -2λ ）は推定値の平滑化を表しています。また、入力 $I_i = d_i - f_i$ は推定値 f_i をデータ d_i に近づける働きをします。

これが、「MRF モデル」です。「結合」はまだついてません。Markov は近傍、Random は確率的、Field は 1 次元または 2 次元の場、という意味です。

3 結合 MRF モデル

さて、このモデルではギザギザしたものは平滑化によってきれいになりました。しかし問題があります。それは画像がもともと持っていたエッジが平滑化によってなくなってしまうということです。理想的なのは、エッジはそのまま保持しつつ、平滑化もする、ということです。これはある意味矛盾してますが、できます。p2.1.m を走らせてみて下さい。

仕掛けは、隠れ変数を導入することによってモデルの自由度を上げる点にあります。これには 2 つのやり方があります。

その1つをまず説明します(テキスト図5)。

$$E(f, l|d) = \frac{1}{2} \sum_i (f_i - d_i)^2 + \frac{\lambda}{2} \sum_i (1 - l_i)(f_{i+1} - f_i)^2 + \theta \sum_i l_i$$

l_i がその隠れ変数です。第1項はデータと推定値は同じでないといけないという条件です。第2項は外界は滑らかであるという条件ですが、先程と違って係数 $(1 - l_i)$ がかかっている。 l_i は0か1をとり、1だと $(f_{i+1} - f_i)^2$ が消えて、 f_i と f_{i+1} の間の平滑化の効果をなくすることができる。別の言い方をすれば、 f_{i+1} と f_i の水平結合を切ることができる。このメカニズムとして shunting inhibition が考えられます。以上のことは、最初にやりたかったことに近いことが分かると思います。何故これができるようになったかという、隠れ変数を入れたからです。しかし l が1ばかりとると平滑化の作用がなくなってしまうので、 l が1をとると E が高くなるように第3項を加えています。 l_i は f_i と f_{i+1} の差が大きいときに1をとるようにしておく。結局、 l と f は双方向的に相互作用していることとなります。つまり、 f は l に、 l は f に依存している。

もう1つのやり方は、隠れ変数として領域のラベルを考えます(テキスト図6)。そして、隣接するラベルが同一のときには平滑化を生かし、違うときには水平結合を切って平滑化が起きないようにする。逆に、隣接する f の値が大きく異なると、対応するラベルが違うものになるように力が働く。

これは先程のモデルと双対な関係にあります(テキスト図7)。エッジが大切というのは両方に言えることです。先程のモデルのように、エッジをダイレクトに表す方法を「境界ベース結合 MRF モデル」といいます。結合というのは、外界を表すプロセスと隠れ状態を表すプロセスが結合しているという意味です。もう1つのモデルでは、エッジを領域ラベルの不連続として暗に表現しています。これを「領域ベース結合 MRF モデル」といいます。

Q: 図ではラベルが2つですが、もっと増やせるのですか？

A: はい。今日は2つある場合しかやってませんが。

Q: l はポテンシャルの絶対値に反応するの？

A: そうです。 l にポラリティーはありません。

4 境界ベースと領域ベースの比較

これから、境界ベースと結合ベースのどちらが脳のモデルとして尤もらしいかを議論します。この議論は最新なので、もしかしたら間違っている可能性もあります。論点は次の2つです。

1. 階層化の容易さ
2. 心理実験との対応

結論から言うと、領域ベースの方が優れている。

4.1 階層化の容易さ

今やりたいことは、受容野が大きく解像度の粗いトップダウンの影響を適応的に切るということです(テキスト図2)。

ラベルプロセスの場合、H.VとL.V.のラベルを照合して、同じであれば切らない、違っていれば切る、ということをするればよい。つまり、localな情報だけで簡単に計算できる。

一方、ラインプロセスの場合、H.V.では粗い解像度でラインプロセスが立っている。L.V.ではよりきめの細かいラインプロセスが存在する。H.V.からの情報はL.V.の広い領域をカバーしている。そして、その情報を切るかどうかを判断するためには、H.V.の細胞がH.V.におけるラインプロセスのどちら側にいるのか、そしてL.V.はL.V.のラインプロセスにおいてどちら側にいるのか、ということを知る必要がある。これはnon-localな計算で、非常に面倒です。

以上のことから、ラベルプロセスの方がラインプロセスに比べて階層化が容易であることが分かります。

Q: これはノイズを除くことを目的としているのですか？

A: ここではそうですが、実際の神経回路網の中にも水平結合や逆行性結合を介したニューロン間の相互作用がある。そういった相互作用を隠れ変数で切るにはどうしたら良いかという一般的な問題を扱っているわけです。

Q: 隠れ変数を使わないといけないのですか？

A: 自由度を余計に入れないと適応的に切るということうまくいかない、というのは事実だと思います。また今は、そのようなモデルに2種類あって、どちらが良いかということを議論しているということです。

4.2 心理実験との対応

(OHPの図で説明)

これはバインディング問題の心理物理実験です。(a)は小窓問題と言われているものです。動いている直線を小窓を通して見ても、対応点がないために垂直方向の情報しか得られない。従って運動ベクトルは一意に求まりません。1次元の情報で2次元の運動ベクトルを決めないといけないので、これは解けない問題なわけです。この心理実験をやると、解はユニークに決まらないはずなんですが、何らかの制約が働いて、垂直方向に動いて見えます。

ここで、(b)のように線分を3つの窓から見るという状況を考えます。この場合、端点の情報に引っ張られて、真ん中の線分の見えが変わります。

線分が短ければ、90度に動いて見える(端点の運動の向き)。長ければ45度に見える(線分の垂直方向の運動の向き)。

また、長さが一定であれば、遮蔽が小さいほど90度に見える。

Q: 運動方向の見えは一定なのか、それとも45度から90度へ回転していくのか？

A: 心理実験でそのようなダイナミクスを計測するのは難しいので、被験者には単に見えの角度で答えてもらう。見えのダイナミクスは刺激の提示時間を変えることによって知ることができる。我々の仮説は 45 度から 90 度へ回転していくというものです。

Q: 実際どのように見えるのか？

A: それは実験的には確かめていません。

Q: 主観的にはどう見えるのか？

A: 我々は biased subject なので、45 度から 90 度へ回転していくように見えます。

(視覚刺激と被験者の応答のビデオを見る)

Q: 被験者は 45 度か 90 度かを答えるのか？

A: そのような 2 件法も、アナログ値を答えるということも、両方やっています。2 つの結果には一貫性があります。

Q: 被験者の視線は？

A: ずっと赤い点を見ています。

(これに対するモデルを OHP で説明)

この図は領域ベースのモデルです。1 次元のモデルなので、領域といっても 1 次元です。境界ベースの場合、境界は 0 次元になります。

知覚としては、最初は 45 度を向いています。ここで、遮蔽部分に非遮蔽部分と同じラベルを貼ろうとするプロセスが自動的に働くものとしています。これによって遮蔽部分の隠れ変数が補完されます。そうすると端点からの情報が流れ込んできて、運動ベクトルの向きが変わっていきます。

線分が長ければ向きの変化が遅いというのは、端点から情報が流れ込むのに時間がかかるということです。

遮蔽部分が多いと遅くなるというのは、ラベルを補完するのに時間がかかる、ということです。

(モデルの挙動をビデオで見る)

遮蔽部分がたくさんあると、運動の向きが 45 度から 90 度へ回転するのに時間がかかっていることが分かります。

(以上で何が言えるか)

心理実験によって速度情報の伝達スピードというものを計算できます。それによると、遮蔽がある場合は非遮蔽の場合より 3.5 倍遅い、ということが分かりました。

この点を、先程の 2 つのモデルで解析してみます (OHP の図)。ここで、情報は端点だけが持っています。直線部分是对应点が存在しないので垂直成分しか運動情報がありません。そこで、水平成分の情報が端点から流れてくる、と考えています。

非遮蔽の場合、いずれのモデルでも線分の長さ L だけに依存します。

遮蔽の場合ですが、境界ベースの場合は、ラインプロセスの値はローカルに決定されるので、遮蔽の大きさ l に依存しないはずですが、一方、領域ベースの場合、遮蔽部分のラベルは見えているところから次々と補完していく必要があるため、 l に依存する。つまり、領域ベースモデルの方だけが L と l の両方に依存する。これは心理実験に合う、ということです。

Q: 非遮蔽の場合というのはただ単に線分を見せているということですか？

A: そうです。

Q: それだと初めから 90 度に動いて見えるような気がするのですが。

A: それは視覚刺激によります。線分の長さなどの実験条件によっては向きが回転していきます。心理実験の結果もあります。

Q: これは水平結合によって実現しているわけですが、上位の階層との相互作用としても説明できませんか？

A: できます。そのどちらでやられているのかということは実験では区別できない。この心理実験で示したかったのは、頭の中で繰り返し計算が行われているのか、ということです。トップダウンでも水平結合でも説明はできます。

Q: 境界ベースモデルでは、つないでいく時間がかからない、というのが良く分からないんですが。

A: ラインプロセスは隣接する強度プロセスだけを見ていて、2 つの値の差が小さければ線を立てないと決めることができます。ラインプロセスも遮蔽の大きさに依存して時間がかかると言われれば、上の議論は終わりです。

Q: それは 1 次元に限って考えているからですか？

A: 2 次元では線同士の相互作用が出てくる、ということはありません。1 次元ではそれがありません。

以上の議論で、階層化の容易さ（計算の局所性）、心理実験との対応、という 2 点から、領域ベースモデルの方が良い、ということが分かりました。

5 位相によるラベルの表現

しかし実はまだ良くない。これをそのまま最急降下法でやると、計算が意味のないところで止まってしまう。というのも、プロセスが非線形になっているので、エネルギー関数に沢山のローカルミニマムが存在して、そこで計算が止まるということが起こるわけです。

それでどうするかというと、アドホックなんですが、ラベルを位相で置き換える。何が違うかというと、ラベルは 1,2,3, というように離散的なんですが、位相は 0 から 2π まで

の周期的な値をとる．ラベルの場合、ラベル1のニューロンが発火するとラベル3のニューロンが発火するというを考えていて、1と3の距離みたいなものはない．位相の場合は周期の中に1とか3とかを割り当てる．そういう表現に置き換えてやります．

このモデルに、計算課題として表面再構成と図と地の分離を行わせる．
モデルは次のようになります．

$$E(f, \phi|d) = \frac{1}{2} \sum_i (f_i - d_i)^2 + \frac{\lambda}{2} \sum_{n.n.} (1 + W_i W_{i'}) (f_i - f_{i'})^2 - \frac{J^R}{2} \sum_{n.n.} W_i W_{i'}$$

第1項はデータフィットの項で、推定値とデータの差をできるだけ小さくする．第2項の $(f_i - f_{i'})^2$ は推定値が隣同士で似ていることを要請する．ここに隠れ状態を入れる． W_i は2次元単位ベクトルですが、角度 ϕ でパラメライズされる．この ϕ を位相と呼ぶ．結局第2項全体によって何が起これるかという、 W_i と $W_{i'}$ が反平行だと内積が -1 になって、第2項が消える．平行だと 1 になって、第2項が効いて平滑化が起これる．

先程の p2.1.m は、実はこの位相をあらかじめ与えた上で表面再構成を行ったものなのです．位相は Figure 2 に出てきたものです．このグラフでは位相を π で割っているので 1 ずれていると逆位相であることを示しています．図の部分と地の部分はそれぞれ同位相、図と地は逆位相になっています．そうすると、それぞれの領域では平滑化が起これますが、境界の部分では起これない．だからうまくいったわけです．これによって、位相が決まればうまくいくということ、つまり隠れ状態の重要性が半分示されたこととなります．

それでは p2.2.m を動かしてください．今度は位相 ϕ を与えずに、 f と ϕ の両方を計算しています．結果は、表面再構成の方は常にうまくいきますが、位相の方は、適切に計算される場合と、半々か、もしかしたらそれより多い確率で、このようになります（エッジを挟んで逆位相になるが、それぞれの面で一定にならず、なだらかに変化している図）．この場合、境界の部分では確かに位相が π 違うので平滑化は起きない．面の部分は、位相が少ししか変わらないので、平滑化は起きる．それで、ある程度うまくいくというわけです．しかしこれはローカルミニマムです．図、地それぞれ一定値を取ったほうが E は小さい．隠れ変数を離散でやるとうまくいかない理由は、このような滑らかな値を取れないところにあります．

結局位相を使ってもうまくいかないのではないか、と言われるかもしれませんが、実はうまくいく（テキスト図 10）．左が位相、右が外界を表す強度で、上の2つが初期状態です．初期状態では図と地について情報が無いので、位相はどの点もランダムにほぼ同じ向きにしています．これを p2.2.m の2次元版で解いてやると、その下の図のようになります．表面再構成はうまく行って、位相も図と地で π 異なります．ここでは位相差だけが問題なので、 2π 飛んでいる部分は位相差 0 ということになります．

ここで、離散にラベルをとると何故よくないかということを説明します．2枚の平面に対して $+1$ か -1 かを振ることにします．ラベルが同じだと $+1$ 、違うと -1 なので、先ほどの W をこれで置き換えてやると離散版になります．これで先ほどと同じタスクを行わせると、（テキスト図 12）ラベルがブランク状に貼られて、図と地の分離がうまくいきません．この場合、ローカルミニマムにトラップされてそれ以上動かない．

このように位相と離散で結果が異なるのは、相互作用の構造の違いが原因です．どちら

のモデルのエネルギー関数にもラベル同士の積があります（テキスト図 13）。離散の場合（ $E = -Js_1s_2$ ）、このようなモデルをイジングスピンといいます。一方、2次元単位ベクトルの位相で表現した場合（ $E = -JW_1W_2 = -J\cos(\phi_1 - \phi_2)$ ）、 E は位相差のみに依存し、その絶対的な値には依存しません。従って J が正の時は $\phi_1 = \phi_2$ に、負の時は差が π の直線上に等エネルギー線が存在します。この線上ではエネルギーの損失なしに動くことができます。これを中立安定と言います。

結局、イジングスピンの場合、点にはまるとそれ以上動けないんですが、位相の場合は状況に応じて等エネルギー線上を無理なく動けるわけです。それで直感的にうまくいきそうな気がします。実際、そのようなプロセスは起こっています（テキスト図 11）。最初は (a) のように図の両端で符号が異なるということが起こり得ます。これは先程見たローカルミニマムの状態です。1次元の場合はこのままで終わるんですが、2次元の場合は1次元の場合とは違って周りからの影響がある。そうすると渦がほどけるようなことが起こって、(b) のようになり、最終的にはうまくいく（テキスト図 10(c)）。結局、位相の場合、無理なく動くことができ、1周すると元に戻る、という性質が効いてうまくいく。

6 議論

ここで、スパイクの同期は情報のキャリアか、という問題を考えてみます。明日青柳先生がお話される位相振動子モデルとここで説明したモデルは対応させることができます。最も単純な位相振動子が今日お話しした ϕ です。従って、スパイクの位相によって ϕ が表現されている、という仮説を立てることができます。

重要なことは、計算理論的に機能で押し込んで、つまり表面再構成に対して何が必要か必要でないかということを考えていて、その結果として位相が必要であるという結論を導いた点です。この位相と位相振動子が結びつければ、もしかしたら随伴現象かも知れないスパイクの同期が、情報のキャリアである、と言えることになります。

発火率がキャリアであることはすでに分かっています。一方で、位相振動子的な要素は神経回路網では簡単に出てくる。物理的に無理はない。しかしそれが情報のキャリアかどうかは分からない。ここで、位相振動子的な要素が計算論的に必要ということであれば、それを使えばいい。これによって、強度プロセスと隠れプロセスに別々のニューロンを割り当てていたところを、1つのニューロンにインプリメントすることが可能になります。

1つのニューロンで行う場合、発火率と位相の相互作用の仕方は同じでないといけない（テキスト図 7）。1つのニューロンで強度プロセスとラインプロセスをやろうとすると、エッジの情報は1次的に、面の情報は2次的に相互作用する必要がある。ラベルプロセスを使った場合には、両方が2次的に相互作用していけば良いので、異なる様式の相互作用を同時にインプリメントするという原理的な困難は存在しない。

また、Singer やそれに触発されて出てきた従来の位相振動子モデルは、外界が似ているとスパイクの同期が高まる、という一方向的なものです。それに対して今述べたモデルは、同位相か逆位相かによって強度プロセスの混ざり具合を変える。強度プロセスと隠れプロセスは双方向に影響し合っている。この点が従来のモデルとは本質的に異なります。

今後の課題としては、1ニューロンでのインプリメントや、実験的に隠れ変数を探すことなどを考えています。また、スパイクの同期が情報のキャリアかどうかを検証する実験を考案することも理論家の仕事だと思っています。以上です。

Q: 従来のモデルというのは振動現象を随伴現象と捉えているのか、または、振動現象を説明するためのモデルということなのか。

A: Singer らはそのようには考えていない。例えば binocular rivalry といった知覚現象とスパイクの同期とは相関がある、つまり neural correlate がある、と言っている。ここではもっと直接的に、スパイクの同期によって知覚が変わる、ということを示しています。

Q: 振動現象の機能的役割を論ずるモデルということですか。

A: そうです。それを計算理論から導いた、ということです。

Q: このモデルで重要なのは、隠れ状態を絶対的な値ではなく相対的な値で表現するということですね？

A: そうです。あと、一周して元に戻ると言うことです。これは先日酒井さんに言われたことですが、同じような力学的構造は、ニューロンを何個か円環状に結んで、メキシカンハットタイプの相互作用を入れてやれば実現できます。こちらでも良いのではないかと、言われれば、否定はできません。問題はインプリメンテーションの容易さだと思います。

Q: 位相の場合、何個までラベルを貼ることができるのか？

A: ラベルを増やすと難しい。ローカルミニマムに入ってしまう。大体3から4個です。

Q: その場合どのようにしているのか？

A: $\cos(\phi_1 - \phi_2)$ を $g(\phi_1 - \phi_2)$ と一般化して、変動を細くする。これによって0から 2π を3個以上に分割する。

Q: そうすると、4個以上だと隣接の関係が変わってくるのではないかと、つまり3個までだとどの状態もそれ以外の状態に隣接している。

A: そうですね。シミュレーションだと4個まではうまくいく。但し成功率は落ちる。

Q: 位相の構造自体を変えるということではできないのか？

A: 3次元単位球の方位を使うとか、

Q: トーラスにするとか、

A: そうですね、そういうことも充分考えられます。その場合、ニューロンが位相方程式に落とせるようなものではなく、もっと豊かさのようなものを持っている必要があります。

- Q: 1 個のニューロンでは無理ということか？
- A: 1 個のニューロンが必ず 1 次元に落とせるというわけではなく、2 次元のモデルもあります。ニューロンが常に位相振動子モデルで記述されるわけではありません。
- Q: ニューロンの豊かさのようなものにラベルの数が依存するということですか？
- A: そういうことです。
- Q: 心理学的には何個ラベルを用意すれば良いと考えられているのか？
- A: よく知らないが、多分 3,4 個だと思う。10 個はできない。奥行き方向の面がいくつ順序立てて知覚できるか、ということになるわけですが。
- Q: 図のほうだけを再構成すれば良いのであって、地の部分を再構成する必要は無いのではないか？ そうだとすれば、相対的な位相を用いる意義はあるのか？
- A: はじめに 2 つを分離する必要はあると思うが、しかし、そのあたりは具体的なものが無いのではっきりとした事は言えません。
- Q: 昨日までの話では、計算論的なモデルを最急降下法で書き下すと Hebb 的な単純な学習則になっていて、生物学的なりアリティがありそうだと、ということでしたが、このモデルもそうなんですか？
- A: 結合は学習できません。
- Q: ニューラルダイナミクスは？
- A: アーキテクチャはそのようなものになります。結合は λ を与えてやらないと決まりませんが、そういったインプリメンテーションはすごく楽にできます。
- Q: 要するに、一番重要な点、新しい点、というのは何ですか？
- A: 純粹にモデルとして新しい点は、隠れ状態に位相を導入すれば、ローカルミニマムから脱出しやすい、ということです。その他は解釈に関してということになります。これまでも位相振動子間の相互作用を変える、というモデルは沢山ある。その相互作用の仕方は従来アドホックに決められていたが、ここではコスト関数から直接導いた。更に、その結果として、双方向的な相互作用が出てきた。また、これまでのモデルは Singer がスパイクの同期現象を見つけたから、それを表現として使う、というものです。ここでは何故スパイクの同期でなければいけないのか、と言う点に関しても答えている。
- Q: 境界ベースと領域ベースで、それぞれ長所短所があると思う。今回は領域ベースが良いと言う事でしたが、双対的な構造を持っているのであれば、短所を補うような形で両方使えば良いのではないか？

- A: 僕もそう思います。ただ、解像度の異なる2つのものが相互作用するときには、領域ベースの方がいい、というのは事実です。他のタスクでは境界ベースの方がいい、ということはある得ると思います。
- Q: 今回の話では領域が連続した場合でしたが、離れた領域を結合するような場合はどのように考えればよいのか？
- A: 今回の心理実験のモデルもある意味離れた領域を結合しているのだから、その線で行けば、どこか共通のセンターを介して引き込んでいくと考えるのが自然だと思う。バインディング問題のセンターはもしかしたら V2 かも知れない。タスクに応じてそういうものがあるのではないかな。
- Q: このモデルでは、違う領域に受容野を持つニューロンの間にも位相が 180 度ずれたところに相関のピークが出るということになるが。
- A: このモデルをもう少し高級なモデルに拡張して、図と地でまったくピークが出ないようにすることは、それほど難しいことではありません。例えば、振動数が違うニューロンをばらまいたときには、同じ面の中にあるニューロンはお互い引き込んでいて、ラグ付きのピークが出る。しかし異なる面のニューロン間では相関が出ない。
- Q: Lamme と Zipser の結果では、図と地で少し遅れて発火率に差が出てくる。この点はどのように説明されるのか？
- A: それは今後の課題ですが、多層にするということです。IT に figure neuron、例えば四角ニューロンみたいなのがあって、その振動が V1 の図の部分の振動とちょうど共鳴するような形で相互作用が起こる、地の部分とは共鳴しないので相互作用が起こらない、そういうふうになると考えています。
- Q: それは、隠れ状態がフィードバックされて V1 の図の領域に影響する、ということですか？
- A: 強度プロセスと隠れプロセスとどちらが先か、という問題ですが、まず実験的な難しさがある。定常状態でないと cross correlation はとれない。JointPSTH もよくわからない。仮にそういう問題が克服されたとしても、2つのプロセスは結合しているので、どちらが先なのかは分離できないかも知れない。
- 先のプログラムでは、位相と強度の時定数は同じではなく、位相の方が千倍早く動けるようにしてあります。これは、コサインの動きにくさを補うということもあるが、強度が決まっていなくて位相が決まらないという構造になっているということかも知れない。しかし、最終的にエッジをきちんと決めるためにはラベルが同じかどうかを知る必要があるのだから、やはり鶏と卵で、双方向的な相互作用というのが重要なんだと思います。

(文責：坂谷 智也、土肥 英三郎)