

神経情報科学サマースクール

Neuro-Informatics Summer School, 2000
(NISS2000)

「脳における予測と文脈生成」

トップダウンの情報はいかにして生成されるのか？

2000年8月8日(火) - 13日(日)

神奈川県葉山町 湘南国際村センター

主催：日本神経回路学会

開校にあたって

校長 塚田 稔（玉川大学，日本神経回路学会会長）

異質なシステムが互いの主体性と協調によって統合される時，そこに新しい情報創発が存在する．若い研究者の柔軟な思考による情報創発がこのチュウトリアル目的の1つである．神経回路の理論的研究と実験的研究の相互の理解のもとに高次脳機能研究に新たな発展を期待する．

プロデューサ 川人 光男（ATR 人間情報通信研究所）

日本に，計算論的神経科学の系統的な教育機関はありませんから，この夏の学校が少しでもその代わりになればと思っています．去年から参加したかったのですが，出張で出れなかったのが，今年はとても楽しみです．若い人と知り合いになって，予測と制御に基づく，高次機能の理解のシナリオについて議論したいです．

ディレクタ 五味 裕章（NTT コミュニケーション科学基礎研究所）

本年はじめてこのスクールのスタッフを仰付かり，体制や役割もよくわからないままに，ファカルティや他のスタッフの方にご迷惑をかけながらようやくテキストができあがりました．しばらくぶりに，このような刺激的で学際的な環境に触れることができるのをとても楽しみにしています．私自身の専門に近い研究では大いに議論し，また少し離れた分野の講義は，私も学生の一人になった気分みなさんと一緒に勉強したいとおもいます．また，みなさん自身の研究のおもしろさも是非紹介してください．1週間という短い期間ですが，みなさんの，そして私自身のよい刺激になることを期待しています．

ディレクタ 阪口 豊（電気通信大学）

昨年に引き続き，夏の学校のスタッフとして参加させていただくことになりました．昨年同様，密度が高く，また自由な雰囲気でのいろいろな議論ができることを楽しみにしています．自分の興味のある分野だけでなく，それ以外の分野に属する講師や参加者と交流を深め，自分の底辺を広げていただければと思います．

ディレクタ 銅谷 賢治（国際電気通信基礎技術研究所）

今年も，熱意あふれる学生の皆さんお会いできるということで，非常に期待しています．今回のテーマは「予測と文脈」で，昨年の「情報表現」をベースに，脳の複雑なダイナミクスを理解するための多角的なアプローチが聞けるものと楽しみにしています．昨年は，私のワンマンディレクター体制で，詰め込みカリキュラムになったきらいがありましたが，今年は，五味，阪口，銅谷の3人ディレクター体制で，スケジュールも参加者間の交流が十分できるよう考慮したつもりです．しかしまだ改善すべき点はいろいろあると思います．今後このスクールをより有意義なものとして育てていくため，参加者の皆さんの積極的な提案と行動を期待しています．

目次

開校にあたって	2
目次	3
初日：プロローグ	4
はじめに：脳における予測と文脈生成（銅谷 賢治）	4
はじめに：脳における双方向情報処理（川人 光男）	5
予測と推定の計算理論的基礎（石井 信，佐藤 雅昭）	14
2日目：ダイナミクス予測 小脳（コーディネータ：片山 正純，北澤 茂）	27
内部モデル学習制御と腕の柔らかさ調節の統合に関して（片山 正純）	28
小脳に学習で獲得される内部モデル（川人 光男）	39
人間の小脳に獲得される内部モデル：脳機能イメージングによる検証（今水 寛）	51
到達運動の最適化と誤差の信号（北澤 茂）	61
運動制御における高次の問題- 到達運動と視覚運動変換を例にとって-（阪口 豊）	73
3日目：報酬の予測 大脳基底核（コーディネータ：小池 康晴，鮫島 和行）	81
強化学習の基礎：Q学習，TD学習（小池 康晴，鮫島 和行）	82
大脳基底核による眼球運動制御：報酬と動機づけの意味（彦坂 興秀）	89
大脳基底核の計算モデル（中原 裕之）	96
4日目：統計的予測 大脳皮質（コーディネータ：岡田 真人，柏野 牧夫）	106
Predictive coding in recurrent neocortical circuits (Rajesh P.N. Rao)	107
Interactions between top-down and bottom-up processing by means of fMRI and psychophysics (Takeo Watanabe)	108
環境に適応し、未来を予測する視知覚システム（西田 真也）	124
聴覚系のダイナミクスと環境適応性（柏野 牧夫）	130
5日目：階層的予測 高次運動野と前頭連合野（コーディネータ：春野 雅彦，池田 思朗）	136
隠れ変数と階層的学習（池田 思朗）	137
予測と制御による階層的運動学習（春野 雅彦）	146
前運動野・一次運動野における運動指令の座標系に見られる階層性について： 空間 座標から筋肉座標へ（寛 慎治）	155
NISS2000 ファカルティ	159

はじめに：脳における予測と文脈生成

銅谷 賢治

(株)国際電気通信基礎技術研究所；科学技術振興事業団 CREST

近年、脳における情報処理は、単に感覚信号を順を追って運動出力に変換していく一方向的、なボトムアップ的なものではなく、きわめてトップダウン的、文脈依存的なものであることが強く意識されている。例えば、サル視覚野の活動記録で、視覚刺激自体は同じでも、どのような課題を行わせているかによって、同じニューロンの活動が異なるといった結果が得られている。また、人間の fMRI 実験などで、感覚入力、運動出力はなるべく同じに揃えた上で、行動課題の違いによる脳活動の変化がさかんに追究されている。

このような、トップダウン的、文脈依存的な脳の振る舞いが、どのような神経回路により実現されているのかということ自体も非常に興味深い問題である。しかしさらに言えば、注意、期待、意図などの情報が、脳の中でどのようにして生成されるのかということが、自律的、目的指向的な脳の情報処理原理を探る上で本質的に重要な問題である。

そこで、今回のスクールでは、「予測」をひとつのキーワードとして、脳の様々な部位での文脈依存的な情報処理のメカニズムを探って見ることにした。近年、計算理論の分野では、ベイズ統計の道具立てを使って、観測された信号から、それを生成するシステムの隠れたパラメータや状態を推定する手法がさかんに研究されている。またそれを、脳の情報処理のモデルとして適用しようという試みも注目を集めている。

例えば、網膜上の 2 次元の視覚刺激から 3 次元の物体の運動を推定する、という場合、3 次元運動に関する現在の推定をもとに視覚刺激を予測し、それを実際の視覚入力と比較することにより、3 次元運動に関する推定値が更新される。このように、文脈依存的な予測と文脈情報の更新が表裏一体の形で進められる情報処理の例は非常に多く、脳に遍在する自己回帰的なダイナミクスを、このような観点からモデル化することは、そのより深い理解につながることを期待できる。

今回、予測の様々な形態として、「ダイナミクスの予測」「報酬の予測」「統計的予測」「階層的予測」という 4 つの計算論的枠組みを考え、それぞれ、小脳、大脳基底核、大脳感覚野、大脳運動連合野の実験的知見との対比を試みることにした。これは、小脳、大脳基底核、大脳皮質がそれぞれ、「教師あり学習」「強化学習」「教師なし学習」に専門化した脳部位であるという仮説[1]に基づいたものである。

各講師の方々には、計算理論あるいは神経生物学の基礎を押さえながらも、最近の話題にどんどん踏み込んでいただくようお願いしてある。参加者の皆さんには、単にそれを吸収するだけでなく、そこで何が未解決の問題として残っているのか、それをどういう手段で攻めて行けばいいのか、主体的に考え討論する場となることを期待している。

参考文献

[1] Doya, K. (1999). What are the computations in the cerebellum, the basal ganglia, and the cerebral cortex. *Neural Networks*, 12, 961-974.

はじめに：脳における双方向情報処理

川人光男

A T R 人間情報通信研究所

科学技術振興事業団、創造科学推進事業、川人学習動態脳プロジェクト

アブストラクト

言語などのヒトの高次認知機能を神経科学の研究対象とするためには、サルなどの動物実験で得られた神経科学のミクロなレベルの知識を、計算理論を媒介として、ヒトを対象にした脳活動の計測データや、言語学などの研究と統合する必要がある。このための新しい計算理論とそれを支持するデータを解説する。我々の提唱するアプローチの対極となる、Chomsky が構築した生成文法研究で大前提とされる仮定に対する批判を行いながら、その全体像を俯瞰する。

はじめに

コミュニケーション、言語、意識などのヒトで著しく発達した高次認知機能が脳科学、神経科学の真面目な研究対象になるとは15年前には予想さえできなかった。当時は、神経科学者がそのような高次機能について語ると、真面目な研究に飽きたと思われるのが落ちだった。しかし状況は大きく変わって、例えば、ヒトと動物のコミュニケーションの違いの理解、非言語的なコミュニケーションの理解、言語の獲得過程の理解などが、日本の脳研究の長期戦略目標にあげられるようになった。これは、主に4つの駆動力によると考えられる。

第1は実験的神経科学の急速な進歩である。脳に関する我々の知識は、場所（例えば脳のどの部位にどのような機能が局在しているかなど）と物質（遺伝子、タンパク質など）については、多重電極や光学システムによる神経活動の計測や、分子生物学の手法など新しい実験技術の導入によって、目覚ましく拡大した。

第2は計算論的神経科学の勃興である。場所と物質に関する知識が積み重ねられるほど、それだけでは脳の機能、さらにいえば情報処理の理解には単純に繋がらないことが、明らかになった。その結果、脳の計算原理、情報表現、アルゴリズムなどを真正面から研究する計算論的研究が進展し、システムレベルの実験神経科学との緊密な協同研究が行われるようになった。まだ歴史は浅いが、脳の特定の機能（視覚、運動制御）や特定の部位（小脳、大脳基底核、大脳皮質視覚野）については素晴らしい成果が得られた^{(1), (2)}。

第3は脳活動を外から、脳を傷つけずに測る手法（非侵襲計測法）の著しい発展である。20年前には、空間的解像度のほとんどない脳波が唯一の手段であったが、脳磁計、PET（陽電子断層撮像法）、fMRI（機能的核磁気共鳴画像法）など新しい方法が次々に開発されてきた。これらを用いて、新しい発見が相次いだ。言語に関する例を3つだけあげよう。ヒト小脳が言語などの高次認知機能で活動することがわかった。2か国語を流暢に喋れるヒトのうち、11歳より若くに学習したヒトでは、2つの言語の表現がブローカ野で入り交じっているのに対して、それより遅くに習得したヒトでは、2つの言語が違う場所に表現されていた。より複雑な構造の文章を理解する時に、ブローカ野とウエルニッケ野の脳活動がより大きくなることがわかった。

第4は、心理学、認知科学、言語学、哲学などと、脳神経科学の境界が、ぼやけながら、より高次へと移動する、研究者の心理と人気のある研究トピックスの動きである。例えば Trends in Cognitive Sciences という最近創刊された認知科学のレビュー誌では、記事の1/3が神経科学、1/3が計算理論に関わる。一昔前には、意志決定、感情、注意、意識などは、神経科学などの自然科学が扱える対象ではなく、人文系学問だけが研究していたことを思いだすのさえ困難になって来た。

このように、神経科学がヒトの知性と心の問題に真正面から取り組む気運が盛り上がりつつあるが、いまなお、それは大変困難である。最大のチャレンジは、ヒトに対しては、電気生理学の単一細胞記録の手法や、トレーサーを用いた神経回路同定の解剖学的手法などが用いられない条件の下で、如何にして心と物質を繋ぐかという点にある。非侵襲脳活動計測の手法は、進歩したとはいえ、上の手法とは較べ様もないほど、得られる情報が限られている。非常に多数のニューロンの活動を空間的、時間的に平均化し、それが2次的にひき起こすだろうと思われる、電場、磁場、血流などの変化を間接的に観測するだけなのであるから。ヒトの知性に関わる、個々のニューロンが、例えば言語課題の遂行中にどのように興奮しているか、それを含む神経回路はどのようになっているか、従って1個のニューロンの情報処理がどのように行われているかを観測することは、すくなくとも近い将来は不可能である。現在では想像もできないような方法が開発されない限りは、未来永劫に渡ってそうになってしまう。

従って、ヒト固有の高次機能の研究は、例えば視覚や運動制御など、ヒトと実験動物で共有される機能にくらべて、著るしく困難になる。このことは次のような研究プログラムに関する思考実験を行えば、明らかであろう。仮に、視覚や、運動制御の研究で、使える手法が、計算理論と、心理学（損傷脳の研究を含む）と非侵襲計測の3つしかないとしよう。その結果、当然我々は、高次の視覚野、高次の運動野の存在も、脊髄や小脳の神経回路も、これらの脳部位でのニューロンの発火パターンも、情報表現も、情報処理も知らないことになる。この条件の下では、我々は脳のなかでどのように視覚や運動制御の情報処理が行われているかについて、何も確からしいことは知り得ない。

高次認知機能解明の新しいアプローチ

言語など『ヒト固有の?』高次認知機能の解明は、この3つの手法に頼るしかないから、神経科学としての理解の見込みはないのであろうか。答えは、必ずしもそうとは限らないことを示すのが、今回の講演の最大の目的である。我々は、ヒトの高次認知機能をになう神経回路や神経計算原理そのものには、ミクロなレベルでは、サルとの不連続がないと考える。そうすると、計算理論主導で、サルの神経生理学的研究を、ヒトを対象にした脳活動非侵襲計測研究、認知科学、心理学、言語研究などと結びつけるアプローチが有効となる。この考えの下に、20以上の研究機関による共同研究（目標達成型脳研究）『ヒトを含む霊長類のコミュニケーションの研究』（研究統括、東倉洋一；研究課題提案者、川人光男 <http://www.atr.co.jp/kbp/index.html>）を行ってきた。これは、過去3年間に、数編のナーチャ-論文が発表されるなど、大きな成果をおさめている^{(3), (4), (5), (6), (7), (8)}。その計算理論とそれを支持する幅広いデータを紹介しよう。

我々の立場とちょうど反対の極に位置するのが Chomsky が確立した生成文法理論による言語研究である^{(9), (10), (11)}。生成文法研究の目標は、言語機能の解明、すなわち人間の言語に関し

て生物学的に決定されている機能の解明である。次の4つが具体的なテーマとしてあげられている⁽¹¹⁾。(1) 文法の知識：成人の文法に関する知識はどのような性質をもっているか。(2) 文法の獲得：(1)の知識はどのように獲得されたか。(3) 文法の使用：言語の理解・産出に(1)の知識はどのように使用されるか。(4) ハードウェアの理解：上記の機能はどのように脳内で物質的に実現されているか。

研究の大前提として、次の4つの仮定がなされる。(A) 言語の知識を脳内の音韻、単語、文などのシンボルによる表象とそれに対する離散的、論理的計算と考える(古ぼけた人工知能の古典的計算主義!?)。(B) 言語機能独立仮定=領域固有性：言語機能はそれが無限個の文を産出したり、名詞句の無限の埋め込みを可能にするなど、離散無限を扱うと言う点で、他の認知システムと際立って異なっており、他の認知システムから独立した心、脳のモジュールをなす(言語器官!?)。(C) 言語機能の生得性：生後与えられる言語に関する経験データは非常に貧弱であるので、生得的な言語獲得装置が存在し、これと生後の経験データとの相関から、脳内に個別言語の文法知識が形成される(文法遺伝子!?)。(D) 言語機能の種固有性：この言語機能は人間という種に固有である(『ヒトに固有の』言語はヒトに固有である!?)。ただし、括弧内は、我々がこれらの仮定を批判的にシンボル化して書いたものであるが、後で説明するように、言語器官、文法遺伝子などは実際に使われている用語である。

生成文法の4つの仮定に関する論争

これら4つの仮定((A) 古典的計算主義、(B) 領域固有性、(C) 生得性、(D) 種固有性)に対しては、過去にも多くの批判があり、今なお活発な論争が続いている。例えば良く知られたところでは、種固有性に関する、霊長類のアイやカンジの言語獲得能力にまつわる論争である^{(10), (12)}。今回の講演では、領域固有性と種固有性に対立する新しい仮説、生得性と領域固有性に対立する新しいニューラルネットワーク、そして古典的計算主義を乗り越える新しい計算理論を紹介する。

これらの論争に関わる最近のトピックスを紹介しながら、なぜ生成文法研究と非侵襲脳活動計測を単純に組み合わせるだけでは、言語の脳研究となりえないのか、そして我々が目指す新しいアプローチが必要となるのかを説明していこう。まず(B) 領域固有性、(C) 生得性、(D) 種固有性すべてにかかわる『文法遺伝子』の研究の最新動向を紹介しよう。Pinker のベストセラー⁽¹³⁾では、ネーチャ-、サイエンスなどにも喧しく宣伝された『文法遺伝子』の研究を、次のように紹介している。KE家と言う家系の成員の約半数に見られる言語障害は、活用語尾の使用に限定された障害で、言語以

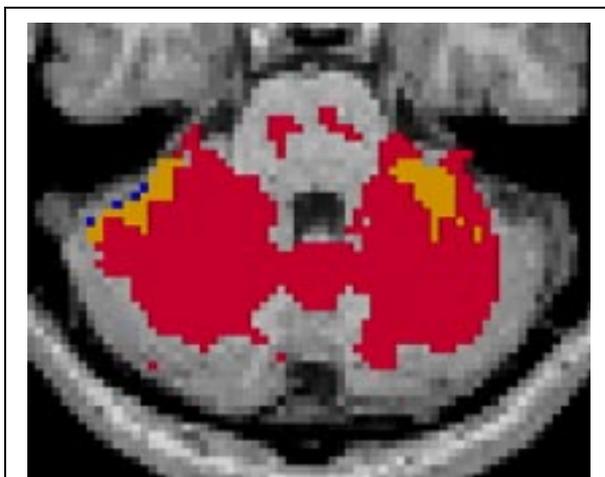


図1 ヒト小脳の道具使用の学習にともなうfMRI活動⁽⁵⁾。図は水平断面を示し、上が前、下が後ろに相当する。被験者は新しい道具の操作方法を学習した。学習初期には赤と黄で示した広い範囲で、道具操作の誤りを表現する信号が計測されたが、十分な学習後は青と黄で示したごく一部に道具の内部モデルが獲得された。これは多重順逆内部モデルの予測と良く一致するデータである。

外の認知能力には異常が見られず、言語のみに障害が見られるいわゆる言語特定性障害である。しかも、系図から、この障害をもたらす遺伝子は、単一の常染色体上にあることが示唆された。これこそ、(B)領域固有性、(C)生得性、(D)種固有性を証明する『文法遺伝子』の証拠であるとマスコミでも大騒ぎされたというのである。ところが最近出版された、より精密で包括的な行動研究、遺伝学研究⁽¹⁴⁾そして非侵襲脳活動計測研究⁽¹⁵⁾は、これとちょうど逆の結論を導いている。障害は、文法のすべての側面に及ぶだけでなく、知能指数など一般的な認知機能、さらには、言語とかかわりのない、顔と口の系列的な運動能力全体に渡っている。PETで観測された脳活動の異常は、前補足運動野、帯状回運動野、補足運動野、感覚運動野、側頭葉、腹側前頭前野、運動前野、尾状核、ブローカ野などなど実に広い範囲に渡っている⁽¹⁵⁾。しかも、構造MRIで発見されたもっとも顕著な脳の構造的な異常は、大脳基底核の尾状核が両側に小さいことだった⁽¹⁵⁾。また、異常がある遺伝子の座は、自閉症に関連する遺伝子がある範囲と一致していた⁽¹⁴⁾。このデータは、文法と言う言語機能を支える遺伝子から神経機構までのハードウェアが、時系列的な感覚運動統合から、コミュニケーションを含む様々な認知機能に関わっているという連続性を強く示唆している。KE家のデータは、(B)領域固有性、(C)生得性、(D)種固有性とはむしろ矛盾していたのである。

我々は『大脳皮質、小脳、大脳基底核がなす閉ループは、時系列的な感覚運動統合から、非言語コミュニケーション、言語までの様々の情報処理を、脳内のミクロなレベルでは同じ神経計算原理に基づいて解いている』という仮説を提案する。最近5年間に、小脳と大脳基底核が高次認知機能(言語、想像、自閉症、分裂病)に関わっていることを示す沢山の非侵襲脳活動計測と臨床的なデータが得られて、小脳と大脳基底核が、運動制御だけに限られる脳部位であるという古い考えはすてられつつある。それにかわって、大脳皮質、小脳、大脳基底核はそれぞれ、異なる学習アルゴリズムを実行していると考え、なぜ感覚運動統合から、言語などの高次認知機能まで、3つの脳部位が役割分担をするのか理解できる⁽²⁾。特に小脳は、系統発生的に新しい部分も古い部分も皮質の神経回路は同じで、シナプス可塑性も場所によらないので、同じ神経計算原理が、運動制御から言語まで働いていることが分かる。小脳の系統発生的に古い部分は、身体など運動制御対象の内部モデルを学習で獲得する⁽¹⁾のに対して、小脳の系統発生的に新しい部分は、道具⁽⁵⁾、他者、他の脳部位のモデルを獲得すると考えられる(図1)。この研究成果によって、小脳の内部モデル仮説は高次認知機能へ拡張されることになった。コミュニケーションで重要な役割を果たす『心の理論』とは、ヒトが他人の心の状態を推定したり、操作したりする時に使うと考えられる、シミュレータ、エミ

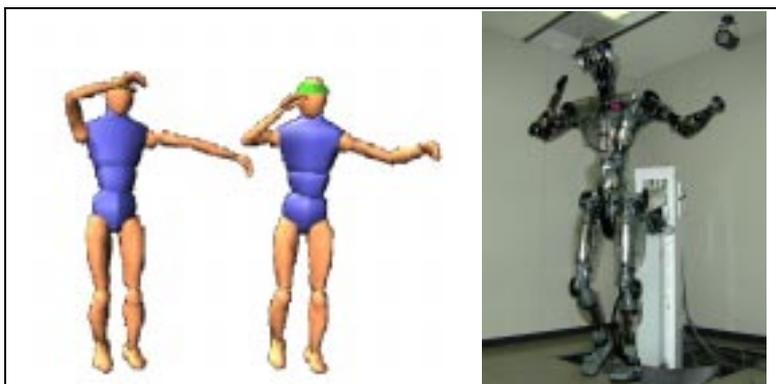


図2 沖縄舞踊カチャーシを見まねで踊る人間型ロボットDB (<http://www.erato.atr.co.jp/DB/>)。ロボットの視覚に相当するシステムで計測された先生の踊り(左)はロボットに可能な動き(中)に変換され、実際に踊られる(右)。見まねはコミュニケーションの最も基本的な素過程であり、様々なレベルでの内部モデルを必要とする。

ユレータ、あるいは他人の心に関する科学の仮説のような『理論』である。小脳内部モデル仮説に従えば、『心の理論』は他者の脳の内部モデルとなる。これが学習で獲得できることを示そう。見まねの能力は、このような様々なレベルでの内部モデルに依存している（図2）

大脳基底核は行動の評価に関わり、将来の行動の方向付けをする⁽²⁾。一方、大脳皮質のブローカ野などを含む前頭葉の役割は、必要な情報を一時的に保持するワーキングメモリであり、これは統計的学習と連想記憶ダイナミクスで実現されている。ヒト知性に関する我々の考えの新しい点は、これまで高次認知機能を前頭前野など大脳皮質だけにおいてきた考えを排し、小脳、大脳基底核との多重で階層的なループこそが、ヒト知性の神経機構の秘密だと考える点に有る。Pinker が、言語器官と呼んだ⁽¹³⁾ ブローカ野、ウエルニッケ野を含むシルビウス溝周辺領野は、決して言語と言う機能に固有のハードウェアではないことが、サルの脳の解剖学と電気生理学の研究からも^{(16), (17), (18)}、ヒトの非侵襲脳活動計測⁽¹⁷⁾からも明らかになってきた。それは、むしろ運動系列を制御し、予測し⁽¹⁹⁾、計画し、認知するための一般的な感覚運動統合のための器官なのであり、それがコミュニケーションから言語へと転用されたと考えられるのである。

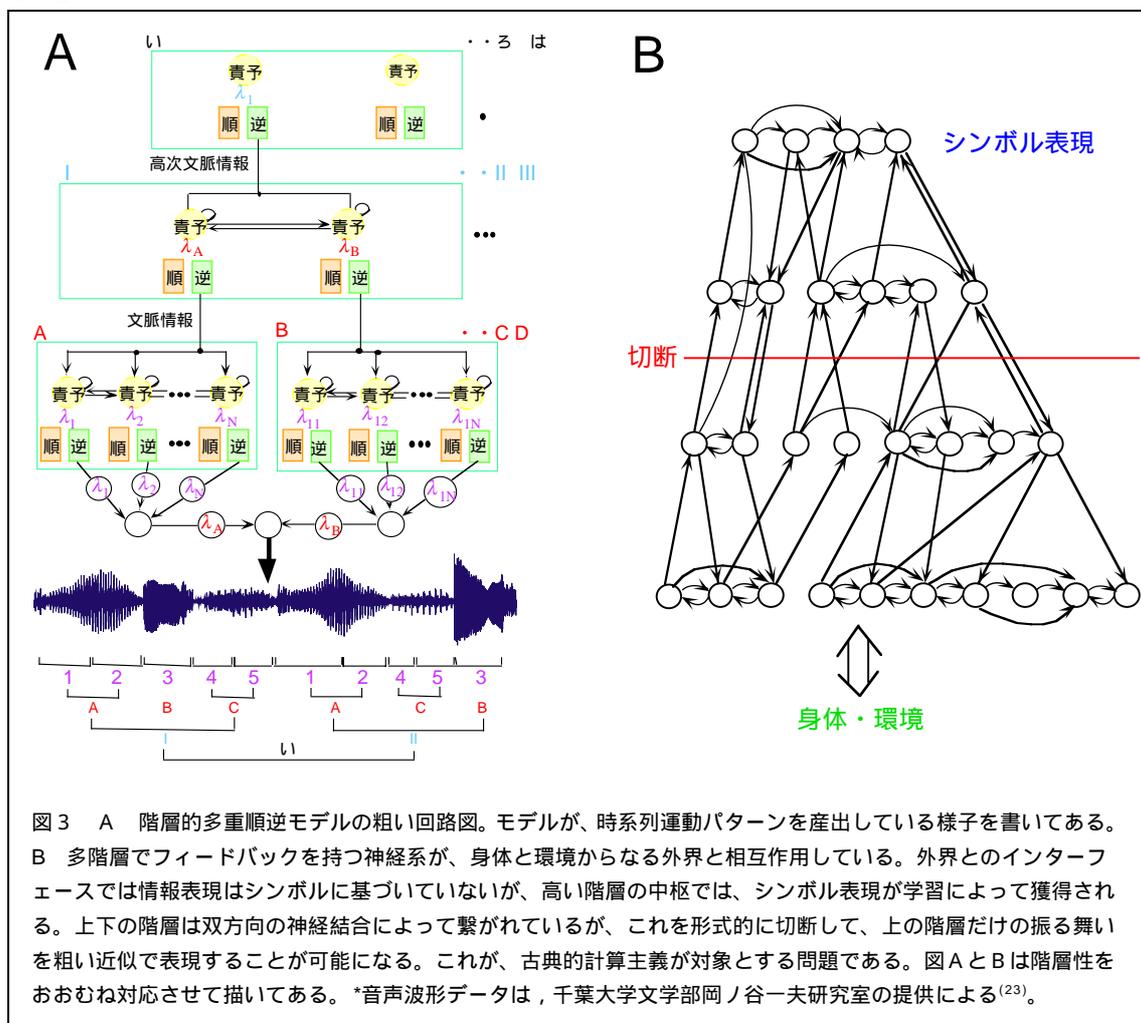


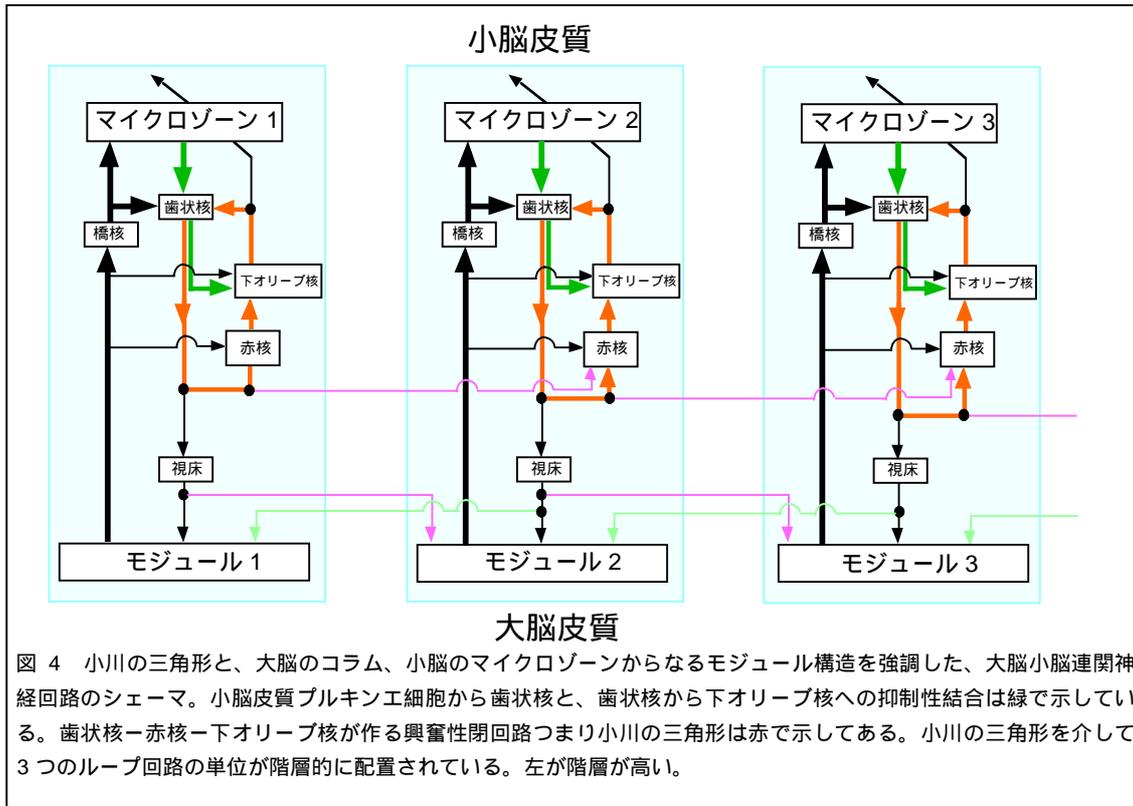
図3 A 階層的な多重順逆モデルの粗い回路図。モデルが、時系列運動パターンを産出している様子を書いてある。B 多階層でフィードバックを持つ神経系が、身体と環境からなる外界と相互作用している。外界とのインターフェースでは情報表現はシンボルに基づいていないが、高い階層の中枢では、シンボル表現が学習によって獲得される。上下の階層は双方向の神経結合によって繋がれているが、これを形式的に切断して、上の階層だけの振る舞いを粗い近似で表現することが可能になる。これが、古典的計算主義が対象とする問題である。図AとBは階層性をおおむね対応させて描いてある。*音声波形データは、千葉大学文学部岡ノ谷一夫研究室の提供による⁽²³⁾。

Chomsky 自身は、生成文法を、原子の構造が発見される以前の化学の周期律表に喩えている。では将来、言語の脳科学がミクロなレベルで大きな進歩を遂げた時に、生成文法は、統計

力学に対する熱力学のように、正しい現象論の役割を果たせるのであろうか。神経科学の例で言えば、視覚や運動制御の計算論的研究で重要な役割を果たした、コンピュータビジョンのアルゴリズム、あるいはロボティクスの理論に相当する役割を期待できるのだろうか。残念ながらそれが甚だ疑わしいのである。最も研究が進んでいる統語論に限っても、生成文法理論はコンピュータによる自然言語処理などの現場で役に立たないのである。正しい構文の可能性を莫大な数生成してしまったり、文法的に誤った文は処理できなかつたりする。生成文法研究の言語学としての成功の最大の理由は、シンボルの離散的・論理的計算に研究の的を絞ったこと（(A) 古典的計算主義の仮定）にある。しかし残念ながら、まさにこれが実際の言語データを情報処理できない本質的理由になっている。つまり、シンボルと論理計算だけを取り扱うと言うことは、シンボルにならない神経表現、シンボルと身体及び環境を含む実世界との対応、意識下で行われる様々な情報処理、論理計算では表現できない力学系、統計性などが、すべて切り捨てられているのである。言語の神経科学を指導する理論に、言語情報処理を実際に行なえることを最低の条件として要求すれば、この欠点は致命的である。

言語獲得の生得性の仮定（C）に関して、Elman らコネクショニストの言語学者の研究が大きな論争をまきおこしている^{(20), (9), (10)}。簡単なニューラルネットワークモデルと言う力学系と統計的な学習を組み合わせるだけで、言語学習の一側面が説明できるので、いわゆる『生得性』の一部を仮定する必要がないというのである。つまりこのモデルに、言語と言う領域に固有ではない、ニューラルネットワークの構造や、シナプス可塑性の規則、ワーキングメモリの容量が成長とともに増加することなど、一般的な制約だけを与えて、大量の文を訓練データとして学習させると、ある語系列を提示された時に、次に出現可能な単語が予測できるようになる。最近の研究では、埋め込み文の深さに対応するカウンターが、学習で力学系の中に獲得できることが示された⁽²¹⁾。この結果は、ニューラルネットワークモデルでは、任意の深さの埋め込みを陽に扱うことはできないとする生成文法研究者からの従来の批判が正しくないことを示している。Elman らのアプローチは、古典的計算主義（A）の離散的・論理的計算ではない、力学系と統計的学習を導入すれば、生得性の仮定（C）が必須ではないことを示した点で高く評価できる。しかし、ニューラルネットワークの構成要素であるニューロンの情報表現に単語と言うシンボルを用いていることから、古典的計算主義（A）の最初の呪縛つまりシンボル表現からは抜け出していないところが、致命的な欠点になっている。

我々が提案するニューラルネットワーク（階層的な多重順逆内部モデル：図3A）は、力学系と統計的学習を計算機構として含み、情報表現としては、環境と接する面では、音韻、単語、文などのシンボル表現を用いないのはもちろん、中枢部においても先験的にはシンボル表現は仮定しない。しかし、学習によって、中枢に階層性とシンボル表現が自動的に獲得されるのである。我々は、シンボルを情報処理の操作対象となりうる神経表現であると定義する。この意味で、単に感覚運動信号を有限個の集合に分節化して、それを代表する表現（例えば自己組織化マップの量子化ベクトルなど）は、それだけではシンボルとはなりえない。階層的な多重順逆内部モデルはもともとは純粋な計算モデルとして提案されたが、小脳と大脳の連関ループ、さらには特にヒトでいちじるしく顕著な、小川の三角形（小脳歯状核、中脳の小細胞性赤核、脳幹の下オリーブ核が構成する興奮性神経結合の閉ループ）を中心とする神経回路と見事に対応する（図4）。



このニューラルネットワークに基づいて、古典的計算主義（A）にとって変わる、しかしさりとて Brooks⁽²²⁾ のような単純な昆虫ロボットにしか使えない相互作用主義（直接知覚論）でもない、新しい計算理論のパラダイムを提案する。言語の問題に即して言えば、言語の神経計算原理を明らかにする時、もっとも本質的な問題は、シンボルが、シンボルのない神経情報表現（例えば、筋肉への運動指令、視覚などの感覚入力）から、どのように系統あるいは個体発生するかと言う点にある。この部分に関してさえ基本的な理解が得られれば、シンボルにならないレベルの情報を切り捨てて、シンボル間だけの法則性を記述すること（統語論）は半ば自動的に理論から派生することになる（図3 B参照）。この時、生成文法理論は、新しい計算理論がシンボルだけのレベルで妥当か否かを検証する試金石の役割を果たせると期待できる。ヒトの知性一般に敷衍すると、シンボルが先験的な表象として存在しない神経回路に、それが身体や外界や他の個体との相互作用をした結果、シンボルが、ハードウェアとして、もしくは力学系のアトラクターの分節化として出現すると言うシナリオである。しかも、下の階層と外界との相互作用を切り捨てて、上位のシンボルが表現されている層だけを観測すれば、一見意味のある規則性が力学系として存在する（図3 B参照）。それは近似的には、離散的・論理的計算で表現できるものになるだろう。この新たな計算理論は、シンボルに基づいていないという意味で、古典的計算主義とは決別し、シンボルを産出すると言う意味で、相互作用主義とも決定的に異なる。

おわりに

言語などヒト固有の知性の脳研究の2つのアプローチをまとめて比較しておこう(表1参照)。第1は、生成文法が代表する、ヒトとそれ以外の動物の間に、ハードウェアから言語現象にい

たるまで、徹底的な不連続を仮定する立場である。第2の、我々が提唱するアプローチは、ヒト知性の生物学的進化を仮定して、ハードウェアのレベルでは、つまり神経機構や神経計算原理としては、質的な不連続はないとする立場である。力学系の分岐現象のように、ハードウェア（遺伝子、たんぱく質、神経回路など）の小さなパラメータの変動によって、結果として、他の動物から不連続に見える並外れたコミュニケーション能力や言語が獲得されたと考えるのである。現時点ではどちらが正しいかの決着はついていない。しかし、第1の立場を取る限り、本稿で述べてきた理由で、ヒト知性を神経科学として理解することはできない。従って、論理的な帰結として、我々は第2の可能性にかけているのである。

核廃棄物の臨界事故を例に取ろう。大事故が起きた後で、その悲惨な結末をいくら詳細に記述しても、臨界現象のメカニズムの本質には迫れない。生成文法の統語論は事故が生じた後の工場の様子を記述することに相当する。また、閾値をこえる最後の一滴の廃棄物には、求める答えはない。文法遺伝子にヒト知性の根源を求めるのは、最後の一滴の廃棄物を詳しく調べるようなものである。一見、的外れのように見えて、臨界以下の現象を精密に調べ、臨界の前後で共通のメカニズムを明らかにすることがもっとも合理的なのである。感覚運動系列の学習や非言語コミュニケーションと言語に共通の原理を調べることがこれに相当する。

	古典的な アプローチ	新しい アプローチ
計算理論	離散的・論理的計算	力学系・統計学習からシンボルへ
情報表現	シンボル	連続表現からシンボルが生成される
領域固有性	固有：言語器官	ハードウェアとしての領域固有性はない
生得性	生得的言語獲得 機構：文法遺伝子	神経機構の大域的構造・可塑性は生得的：後は学習
種固有性	固有	力学系の分岐現象のように現象としては不連続だが機構・パラメータは連続
何を明らかにしたいのか	現象論	神経計算原理

表1 ヒト高次認知機能の研究で可能となる2つのアプローチの比較

参考文献

- (1) 川人光男：脳の計算理論，産業図書（1996）
- (2) 銅谷賢治：運動学習の神経計算機構 - 基底核，小脳と大脳皮質，別冊・数理科学「脳科学の frontline」，141-152（1997）
- (3) Kitazawa S, Kimura T, Yin P: Cerebellar complex spikes encode both destinations and errors in arm movements. Nature, 392, 494-497（1998）
- (4) Sugase Y, Yamane S, Kawano K, Ueno S: Global and fine information coded by single neurons in the temporal visual cortex. Nature, 400, 869-873（1999）
- (5) Imamizu H, Miyauchi S, Tamada T, Sasaki Y, Takino R, P tz B, Yoshioka T, Kawato M: Human cerebellar activity reflecting an acquired internal model of a novel tool. Nature, 403, 192-195（2000）
- (6) Inui T, Otsu Y, Tanaka S, Okada T, Nishizawa S, Konishi J: A functional MRI analysis of comprehension processes of Japanese sentences. NeuroReport, 9, 3325-3328（1998）

- (7) Wolpert D, Kawato M: Multiple paired forward and inverse models for motor control. *Neural Networks*, 11, 1317-1329 (1998)
- (8) Perrett DI, Lee KJ, Penton-Voak I, Rowland D, Yoshikawa S, Burt DM, Henzi SP, Castles DL, Akamatsu S: Effects of sexual dimorphism on facial attractiveness. *Nature*, 394, 884-887 (1998)
- (9) 伊藤正男他編：岩波講座認知科学 7, 岩波書店 (1995)
- (10) 大津由紀雄他編：岩波講座言語の科学 10 言語の獲得と喪失, 岩波書店 (1999)
- (11) 郡司隆男：情報数学セミナー 自然言語, 日本評論社 (1994)
- (12) 松沢哲郎：チンパンジー・マインド：心と認識の世界, 岩波書店 (1991)
- (13) Pinker S: *The Language Instinct*, Harperperennial Library (1994); 日本語版, 言語を生み出す本能, 椋田直子訳, 日本放送出版協会 (1995)
- (14) Fisher SE, Vargha-Khadem F, Watkins KE, Monaco AP, Pembrey ME: Localisation of a gene implicated in a severe speech and language disorder. *Nature Genetics*, 18, 168-170 (1998)
- (15) Vargha-Khadem F, Watkins KE, Price CJ, Ashburner J, Alcock KJ, Connelly A, Frackowiak RS, Friston KJ, Pembrey ME, Mishkin M, Gadian DG, Passingham RE: Neural basis of an inherited speech and language disorder. *Proc. Natl. Acad. Sci. USA*, 95, 12695-12700 (1998)
- (16) Petrides M, Pandya DN: Comparative architectonic analysis of the human and the macaque frontal cortex. in "Handbook of Neuropsychology 9" Boller F, Grafman J (Eds), Elsevier, pp.17-58 (1994)
- (17) Rizzolatti G, Arbib MA: Language within our grasp. *Trends in Neurosci.*, 21, 188-194 (1998)
- (18) 小嶋祥三：ミラーニューロンと言語の起源. *岩波科学*, 69, 404-408 (1999)
- (19) 乾敏郎：運動系列予測学習仮説. *神経心理学*, 14, 144-149 (1998)
- (20) Elman JL, Bates EA, Johnson MH, Karmiloff-Smith A, Parisi D, Plunkett K: *Rethinking Innateness*, MIT Press (1996); 日本語版, 認知発達と生得性, 乾敏郎他訳, 共立出版 (1998)
- (21) Rodriguez P, Wiles J, Elman JL: A recurrent neural network that learns to count. *Connection Science*, 11, 5-40 (1999)
- (22) Brooks RA: Intelligence without representation. *Artificial Intelligence*, 47, 139-160 (1991), 日本語版, 表象なしの知能, 柴田正良訳, 現代思想, 18, 85-105 (1990)
- (23) 星野力, 銅谷賢治：生物における系列学習運動に関する考察. 電子情報通信学会技術研究報告, NC99-108, 117-124 (2000)

予測と推定の計算理論的基礎

石井 信 (奈良先端科学技術大学院大学)

佐藤 雅昭 (国際電気通信基礎技術研究所)

1 はじめに

我々がある状況において行動する過程について考えてみる。時刻 t における状況を $x(t)$ 、その時の行動を $y(t)$ とすると、行動の選択過程は $y(t) = F(x(t))$ と表現することができる。脳の行なっていることの多くを高度に抽象化すれば、この F を決めていることに他ならない。例えば人間の運動の場合、 $x(t)$ は環境および身体の状態全てを含むものであり、 $y(t)$ は筋肉を制御する信号である。

何らかの経験によって F を決定することを「学習」と総称する。各 $x(t)$ について模範となるべき $y(t)$ が経験を通じて教えられる時、教師あり学習という。あるいは、関数 F を決めるという意味で、関数近似とも呼ばれる。本稿では、この関数近似の問題について述べる。

脳において関数近似ができたとして、その重要な機能は予測である。我々は日常の多くの場面で予測を行なっている。例えば自動車の動きが予測できなければ道路も横断できない。話相手の反応を予測できなければスムーズなコミュニケーションもできないであろう。こういう予測は時系列の予測として抽象化できる。

予測のもう一つの重要な問題は内部状態の予測である。我々が大きな自由度を持つシステムの挙動の予測を行なうことができるのは、その内部状態に関する予測を行なっているからである。例えば新聞に目を通して、政治や経済の情勢から株式投資銘柄を決めることができるのは、政治経済システムの状態をおぼろげながらも予測しているからに他ならない。コミュニケーションにおいても、相手との会話を通じて相手の内部状態の予測を行ない、予測に基づき良い行動の選択を行なっていると考えられる。

脳の計算理論を考える上では、関数近似と予測は重要な問題である。そこで、本稿では、その理論的な基礎を説明する。

2 最小二乗法と関数近似

入力変数 x と出力変数 y との間の関数関係を観測データから推定するのが関数近似の問題である。その際にはいかなる関数 (族) を用いて近似するのかをア priori に与えることがしばしば行なわれる。この場合、関数族をモデル、関数族から一つの関数を選び出すことをパラメータ推定と呼ぶ。

最も簡単な場合として、 T 個の観測データからなるデータセット $\{(x(t), y(t)) | t = 1, \dots, T\}$ に対して、線形モデル $y = ax$ を仮定する。 x および y は 1 次元 (スカラー) とする。スカラー a を決めるのがパラメータ推定である。その推定の規範として、二乗誤差

$$E = \sum_{t=1}^T (y(t) - ax(t))^2 \quad (1)$$

を最小にすることを考える。二乗誤差関数 E はデータセットが与えられた際のパラメータ a についての関数である。最小点では停留条件が成り立つので、

$$\frac{\partial E}{\partial a} = -2 \sum_{t=1}^T (y(t) - ax(t))x(t) = 2 \sum_{t=1}^T (ax^2(t) - x(t)y(t)) = 0 \quad (2)$$

となる。すなわち必要条件 (十分条件でもある) は、

$$a = \frac{\sum_{t=1}^T x(t)y(t)}{\sum_{t=1}^T x^2(t)} \quad (3)$$

である。(3) 式の分子は入出力データの相関、分母は入力データの分散という形をしている。これを最小二乗法という。ここでの議論は 1 次元で行なったが、スカラー同士の積を内積に替えることにより、ベクトル (多次元) の場合も全く同様になる。

しかし $y = ax$ は原点を通る直線であり、それによって近似できる関数は極めて限られる。そこで M 個の関数のセット $\{k_i(x) | i = 1, \dots, M\}$ を用いて、その線形結合により関数を近似することを考える。

$$y = \sum_{j=1}^M a_j k_j(x) \quad (4)$$

ここで各 $k_i(x)$ はカーネル関数と呼ばれる。カーネル関数は一般に非線形なものが用いられるが、パラメータセット $\{a_i | i = 1, \dots, M\}$ から見ると線形モデルであるため、パラメータ線形モデルと呼ばれる。パラメータ線形モデルの例をいくつか挙げておく。

- 多項式近似

$$y = a_0 + a_1x + a_2x^2 + \dots + a_{M-1}x^{M-1} = \sum_{j=0}^{M-1} a_j x^j \quad (5)$$

- 動径基底関数 (Radial basis functions; RBF)

$$y = \sum_{j=1}^M a_j \exp\left(-\frac{1}{2\sigma^2}|x - \mu_j|^2\right) \quad (6)$$

ここで $|\cdot|$ はユークリッドノルムである。

近年盛んに研究されているサポートベクタマシン (Support vector machine; SVM) もパラメータ線形モデルの一つであるが、最小二乗法と若干異なる規範でパラメータ推定が行なわれる。

データセット $\{(x(t), y(t)) | t = 1, \dots, T\}$ に対する二乗誤差

$$E = \sum_{t=1}^T \left| y(t) - \sum_{j=1}^M a_j k_j(x(t)) \right|^2 \quad (7)$$

による最小二乗法を用いてパラメータを決める。上と同様に、

$$\begin{aligned} \frac{\partial E}{\partial a_i} &= -2 \sum_{t=1}^T \left(y(t) - \sum_{j=1}^M a_j k_j(x(t)) \right) k_i(x(t)) \\ &= 2 \sum_{t=1}^T k_i(x(t)) \sum_{j=1}^M a_j k_j(x(t)) - 2 \sum_{t=1}^T k_i(x(t)) y(t) = 0 \end{aligned} \quad (8)$$

となる。

$$z_i \equiv \frac{1}{T} \sum_{t=1}^T k_i(x(t))y(t) \quad (9a)$$

$$K_{i,j} \equiv \frac{1}{T} \sum_{t=1}^T k_i(x(t))k_j(x(t)) \quad (9b)$$

を定義することにより、必要条件 (8) は

$$\sum_{j=1}^M K_{i,j}a_j = z_i \quad (i = 1, \dots, M) \quad (10)$$

と M 次の連立方程式、あるいは、ベクトル表記で

$$Ka = z \quad (11)$$

と書くことができる。カーネル共分散行列 K は対称 $K_{i,j} = K_{j,i}$ であり、かつ任意のベクトル v に対して、 $v'Kv = \frac{1}{T} \sum_{t=1}^T \left(\sum_{i=1}^M k_i(x(t))v_i \right)^2 \geq 0$ であるから、非負行列である。ここでプライム ($'$) は転置を表す。また特別な場合を除いて K は正定値行列であるので、逆行列が存在して、

$$a = K^{-1}z \quad (12)$$

となる。これがパラメータ線形モデルに対する最小二乗法である。また K の正定値性は、最小二乗解の十分性を意味している。

パラメータに関して非線形なモデルも用いられる。最も有名なモデルが階層型パーセプトロン (Multi-layered perceptron; MLP) である。今までの記述との対応を考慮した例として、以下の 3 層の MLP を考える。

$$y = \sum_{i=1}^M a_i \text{sig}(h_i) \quad (13a)$$

$$h_i = \sum_{j=1}^N b_{i,j}x_j \quad (13b)$$

N は入力次元、 M はカーネル (中間層ユニット) の数である。関数 sig は非線形で有界な関数なら何でも良いが、一般にはシグモイド関数

$$\text{sig}(h) \equiv \frac{1}{1 + \exp(-h)} \quad (14)$$

が用いられる。MLP では 2 種のパラメータがある。中間層から出力への重み a_i については、上記の最小二乗法で行なうこともできるが、一般には $a_i, b_{i,j}$ の学習ともに、勾配法

$$\Delta\theta_k = -\eta \frac{\partial E}{\partial \theta_k} \quad (15)$$

が用いられる。ここで E は二乗誤差関数であり、パラメータである $a_i, b_{i,j}$ についての関数である。 θ_k は a_i と $b_{i,j}$ のいずれかを表現している。また η は正数であり、学習係数と呼ばれる。

勾配法がうまく行く原理を簡単に述べる。パラメータ θ_k が $\Delta\theta_k$ だけ変更される際の二乗誤差関数 E の変化 ΔE について考える。

$$\Delta E \approx \sum_k \Delta\theta_k \frac{\partial E}{\partial \theta_k} = -\eta \sum_k \left(\frac{\partial E}{\partial \theta_k} \right)^2 \leq 0 \quad (16)$$

となるので、二乗誤差関数は (ほぼ) 減少することが分かる。二乗誤差関数には明らかな下界 (0) が存在するので、少なくともそれ以上の値をとる極小点で収束する。すなわち最小二乗推定が実現できる。ただし、パラメータ線形の場合と異なり、大域的な最適性は保証できず、局所的な最適性である。なお、一般に MLP モデルの学習法は誤差逆伝搬法と呼ばれるが、勾配法による (局所) 最小二乗法に他ならない。実際の複雑なアルゴリズムは本質的でなく、その原理は (15) 式で十分である。

(演習 1) データセットおよび基底の中心位置が与えられた場合の動径基底関数のパラメータ推定を最小二乗法を用いて行なう。

$$y = \sum_{j=1}^M a_j \exp\left(-\frac{1}{2\sigma^2}|x - \mu_j|^2\right) \quad (17)$$

基底の中心位置は、しばしば最尤推定法 (後述) で決定される。

3 ダイナミクスの学習

観測変数 x の時系列 $\{x(t)|t = 1, \dots, T\}$ から、その時系列を発生したダイナミクスを推定し、時系列の将来を予測することは重要な問題である。時刻 $t, t-1, \dots, t-M+1$ における観測値の線形結合により、次の時刻での観測変数の推定値 $\hat{x}(t+1)$ を推定するモデルを自己回帰 (Auto regressive; AR) モデルと呼ぶ。

$$\hat{x}(t+1) = a_1x(t) + \dots + a_Mx(t-M+1) \quad (18)$$

M は AR の次数と呼ばれる。

時刻 t について、入力 $X(t) \equiv (x(t), x(t-1), \dots, x(t-M+1))$ 、出力 $y(t) \equiv \hat{x}(t+1)$ の関数近似問題とみなすと、パラメータ線形モデル

$$y(t) = \sum_{i=1}^M a_i X_i \quad (19)$$

による推定問題である。したがって二乗誤差関数

$$E = \sum_{t=M}^{T-1} (x(t+1) - a'X(t))^2 \quad (20)$$

を定義することにより、前章と同じように最小二乗法によりパラメータベクトル a を定めることができる。最小二乗法のための線形方程式 (11) を AR では特に Yule-Walker 方程式と呼ぶこともある。

さて (19) 式の右辺は状態 $X(t)$ の関数であり、

$$y(t) \equiv \hat{x}(t+1) = F(X(t)) \quad (21)$$

の形をしている。これを以後、モデルの時間発展方程式と呼ぶことにする。AR では F として特に線形関数を考えた。一方で、 F として例えば MLP などの非線形関数を考えることもできる。学習、すなわちパラメータの推定の目的は、モデルの時間発展がシステムのダイナミクスに近くなるようにすることである。

データセットを用いてパラメータを決定した後の時間発展方程式、すなわち模倣されたダイナミクスを用いて、時系列の推定を行なうことができる。ARでは(18)式である。これを1ステップ予測と呼ぶ。推定された $\hat{x}(t+1)$ を用いると、 $\hat{X}(t+1) = (\hat{x}(t+1), x(t), \dots, x(t-M+2))$ が得られる。これと関数 F から $\hat{x}(t+2)$ が得られ、それから $\hat{X}(t+2)$ が得られる。こうして次々と時系列の推定ができる。こうした逐次的な推定をマルチステップ予測と呼ぶ。時系列を生成しているシステムがカオスなどの非線形性の強いシステムである場合ARモデルは適当ではないことが多い。その場合、MLPやRBFなどの非線形モデルを用いて推定を行なう必要がある。

システムに外部入力 $u(t)$ がある場合は、(18)式を拡張して、

$$\hat{x}(t+1) = a_1x(t) + \dots + a_Mx(t-M+1) + b_1u(t) + \dots + b_Ku(t-K+1) \quad (22)$$

で推定する手法がある。これをARMA(Auto regressive moving average)と呼ぶ。 K はMAの次数と呼ばれる。ARの場合と同様にパラメータ a, b は $x(t)$ と $u(t)$ の観測データを用いて最小二乗法で決めることができる。

さてこれまでは、暗黙に、システムのダイナミクスを規定する状態変数が観測変数と一致するという状況を仮定していた。したがってモデルの時間発展方程式である(21)式はシステムの状態変数を用いて記述されていることになる。ここでより一般的な場合を考える。システムの状態変数を $Z(t) \equiv (z_1(t), \dots, z_L(t))$ として、システムのダイナミクスを

$$Z(t+1) = F^*(Z(t)) \quad (23)$$

とする。 L をシステム次元と呼ぶ。簡単のためノイズ(システムノイズ)は考えない。観測変数 $X(t)$ は $X(t) \equiv (x_1, \dots, x_N) = CZ(t)$ で与えられるものとする。 N を観測次元と呼ぶ。行列 C は $N \times L$ の観測行列である。

ここで特に考える状況は、観測次元がシステム次元よりも小さい、すなわち $N < L$ の場合である。これを部分観測という。例えば $X(t) = z_1(t)$ とすれば、状態変数の一つだけが観測できることになる。こうした場合でも、システムのダイナミクスである $F^*(Z)$ を、線形モデル、あるいはMLPやRBFなどの非線形モデルによって近似することができる。学習には例えば最小二乗法を用いることができる。この時、システムの状態変数 $Z(t)$ は観測できない変数ということで、隠れ変数、あるいは内部状態変数と呼ばれる。システムが隠れ変数を持つ場合の学習法の多くでは、パラメータの推定と同時に、隠れ変数の推定も行なう必要がある。決定論的な時間発展方程式によりモデル化する場合は、 $Z(t)$ の初期状態が与えられれば、モデルを用いて $Z(t)$ の推定時系列が計算できるので、誤差関数のパラメータによる微分は時間逆向きの誤差逆伝搬法などを用いて計算できる。一方、4章で述べる確率モデルでは隠れ変数 $Z(t)$ がしたがう確率分布を観測時系列 $\{X(t)|t=1, \dots, T\}$ から推定する必要がある。

しかし、隠れ変数の推定を含むダイナミクスの推定は、システムの状態変数が全て観測できる場合に比べて複雑で学習にも多大な計算時間を要する。一方で、埋め込み法を用いることにより、観測時系列から、隠れ変数を推定することなしに、システムのダイナミクスをモデル化することができる。例として状態変数 $Z(t)$ の一つの成分 $z_1(t)$ のみが観測できる場合を考える。時系列 $\{z_1(t)|t=1, \dots\}$ から遅れ座標 $Y(t)$ を以下のように定義する。

$$Y(t) \equiv (z_1(t), z_1(t-\tau), \dots, z_1(t-\tau(M-1))) \quad (t = \tau(M-1) + 1, \dots) \quad (24)$$

τ は遅れ時間、 M は埋め込み次元と呼ばれる。 $M \gg 2L + 1$ が成り立つ時、状態変数 $Z(t)$ で記述されたシステムのダイナミクスを、遅れ座標 $Y(t)$ を用いたダイナミクスに変換できることが示されている。そこで遅れ座標を用いた時間発展則

$$z_1(t+1) = F(Y(t)) \quad (25)$$

でシステムのダイナミクスをモデル化することができる。(18) 式あるいは (21) 式は、(24)(25) 式において $\tau = 1$ とした場合になっている。

(演習 2) 時系列 $\{x(t) | t = 1, \dots, T\}$ に対して AR 法を適用せよ。この時系列は実際に AR モデルによって生成されたものである。次数を色々変えて実験してみよ。マルチステップ予測を行ない、実際の系列と比較せよ。

4 最尤推定法と EM アルゴリズム

確率変数 x の二つの確率分布 $a(x), b(x)$ について、Kullback-Leibler (KL) ダイバージェンス $D(a||b)$ を以下で定義する。

$$D(a||b) \equiv \int dx a(x) \log \left(\frac{a(x)}{b(x)} \right) \quad (26)$$

KL ダイバージェンスは非負であり、0 となるのは (ほとんど全ての x について) $a(x) = b(x)$ である時に限る。これは以下で証明できる。 $z > 0$ に対して、

$$z - \log z \geq 1 \quad (\text{等号は } z = 1 \text{ の時}) \quad (27)$$

であるので、

$$\frac{b(x)}{a(x)} - \log \left(\frac{b(x)}{a(x)} \right) \geq 1 \quad (28)$$

が成り立つ。したがって、

$$\int dx a(x) \left(\frac{b(x)}{a(x)} - \log \left(\frac{b(x)}{a(x)} \right) \right) \geq \int dx a(x) = 1 \quad (29)$$

$a(x)$ が x の確率分布であるという性質を用いた。(29) 式の左辺は $\int dx b(x) + D(a||b) = D(a||b) + 1$ であるから

$$D(a||b) \geq 0 \quad (\text{等号は、ほとんど全ての } x \text{ について } a(x) = b(x) \text{ の時}) \quad (30)$$

となる。以上の性質により KL ダイバージェンスは二つの確率分布間の距離 (差) を計るのにしばしば用いられる。ただし $D(a||b) \neq D(b||a)$ と距離公理のうち対称性を満たしていない。そのため偽距離と呼ばれることもある。

確率変数 x に関する未知のデータ分布 $\rho(x)$ を θ によるパラメータ族 $P(x|\theta)$ で近似することを考える。KL ダイバージェンスは

$$\begin{aligned} D(\rho||P) &= \int dx \rho(x) \log \left(\frac{\rho(x)}{P(x|\theta)} \right) \\ &= \int dx \rho(x) \log \rho(x) - \int dx \rho(x) \log P(x|\theta) \end{aligned} \quad (31)$$

で与えられる。KL ダイバージェンスを最小にすることが近似の目標であり、その時のパラメータ θ が最適と考える。(31) 式右辺第 1 項はパラメータ θ によらない定数であるから、KL ダイバージェンスの最小化は

右辺第 2 項の最大化と等価である。ここで右辺第 2 項は、 $\log P(x|\theta)$ の未知分布 $\rho(x)$ についての期待値である点に注意する。さて、実際のデータ分布が分からないので、右辺第 2 項の期待値の計算は厳密にはできない。しかし、データ数がある程度あれば期待値はデータ平均で近似できる。すなわち、未知の分布 $\rho(x)$ からデータセット $\{x(t)|t = 1, \dots, T\}$ が発生した時、次式が成り立つ。

$$\int dx \rho(x) \log P(x|\theta) \approx \frac{1}{T} \sum_{t=1}^T \log P(x(t)|\theta) \quad (32)$$

右辺は左辺の不偏推定量である。以上のことから、

$$L(\theta) \equiv \sum_{t=1}^T \log P(x(t)|\theta) \quad (33)$$

の最大化を θ について行えば良い。これが最尤推定法である。パラメータ θ についての関数 (33) を対数尤度関数、最尤推定法によって求められたパラメータを最尤推定量という。

さて、入力変数 x が与えられた時に、出力変数 y (D 次元とする) が $y = ax + u$ で与えられると仮定して、データセット $\{(x(t), y(t))|t = 1, \dots, T\}$ から、パラメータ a を求める問題を考える。ただし u は分散 σ^2 のガウスノイズである。このとき、 y は中心位置 ax 、分散 σ^2 の正規分布 $\mathcal{N}(y|ax, \sigma^2)$ にしたがう。

$$P(y|x, a) = \mathcal{N}(y|ax, \sigma^2) \equiv (2\pi\sigma^2)^{-\frac{D}{2}} \exp\left(-\frac{1}{2\sigma^2}(y - ax)^2\right) \quad (34)$$

ここで $\mathcal{N}(z|\mu, \sigma^2)$ は中心 μ 、分散 σ^2 の確率変数 z についての正規分布である。中心位置が入力 x に依存しているため、(34) 式で与えられる y の分布は x の条件付きである点に注意する。(34) 式より、

$$\log P(y|x, a) = -\frac{1}{2\sigma^2}(y - ax)^2 - \frac{D}{2} \log(2\pi\sigma^2) \quad (35)$$

であるので、対数尤度関数は

$$L(a) = -\frac{1}{2\sigma^2} \sum_{t=1}^T (y(t) - ax(t))^2 + (a \text{ に依存しない項}) \quad (36)$$

となる。 $L(a)$ を a について最大化することと、(1) 式の E を a について最小化することは等価である。したがってこの場合の最尤推定法は最小二乗法と等価である。これは最も簡単な場合について述べたが、2 章で述べたような一般的な関数近似モデルに関する最小二乗法も最尤推定法から導くことができる。モデルにガウスノイズが付加された形の確率モデルを考えると、その確率分布は $\exp(-\frac{E}{2\sigma^2})$ の形をしている。ここで E は二乗誤差関数である。すなわち、この場合にも最尤推定法は最小二乗法と等価になる。

最尤推定法を用いるためにはデータの分布に関する確率分布族をアприオリに定める必要がある。最小二乗法の場合と同様にこの分布族をモデルと呼ぶが、特に確率分布のモデルであるので確率モデルと呼ばれる。また、データの出現の仕方をモデル化しているので (データ) 生成モデルと呼ばれることもある。

データの分布を近似することの応用の一つがクラスタリングである。今、 M 個のクラスタの各々が等方分散の正規分布をなして、そこからデータ $\{x(t)|t = 1, \dots, T\}$ が観測されたと仮定する。以下の生成モデルを仮定する。

$$P(i|\theta) = \nu_i \quad (37a)$$

$$P(x|i, \theta) = \mathcal{N}(x|\mu_i, \sigma_i^2) \quad (37b)$$

パラメータは $\theta \equiv \{\nu_i, \mu_i, \sigma_i^2 | i = 1, \dots, M\}$ である。 ν_i は i 番目のクラスからデータが出てきた確率を表し、混合比と呼ばれる。 $P(i|\theta)$ が確率であるために、 $\sum_{i=1}^M \nu_i = 1$ を満たす必要がある。この式から、

$$P(x|\theta) = \sum_{i=1}^M P(x, i|\theta) = \sum_{i=1}^M P(x|i, \theta)P(i|\theta) = \sum_{i=1}^M \nu_i \mathcal{N}(x|\mu_i, \sigma_i^2) \quad (38)$$

となる。データが複数の正規分布の混合から生成されたことを仮定しているため、これを混合正規分布と呼ぶ。

注意すべき点は、混合正規分布 (38) において、クラスターの指標 i は隠れ変数になっていることである。すなわち、データを観測した点では、(37) 式に対応する生成過程は観測できないため、データがどのクラスターから出てきたかは厳密には分からない。できるのは推測することだけである。

一般に、隠れ変数 z を持つ確率モデル $P(x|\theta) = \sum_z P(x, z|\theta)$ の最尤推定法について考える。説明の都合上 z を離散変数とするが、連続の場合でも同様である。データセット $\{x(t) | t = 1, \dots, T\}$ に対する対数尤度

$$L(\theta) \equiv \sum_{t=1}^T \log P(x(t)|\theta) = \sum_{t=1}^T \log \left(\sum_z P(x(t), z|\theta) \right) \quad (39)$$

を最大化する θ を求めたい。必要条件である停留条件 $\partial L / \partial \theta = 0$ は以下のように計算できる。

$$\begin{aligned} \frac{\partial L(\theta)}{\partial \theta} &= \sum_{t=1}^T \sum_z \frac{\partial P(x(t), z|\theta) / \partial \theta}{\sum_{z'} P(x(t), z'|\theta)} \\ &= \sum_{t=1}^T \sum_z P(z|x(t), \theta) \frac{\partial}{\partial \theta} \log P(x(t), z|\theta) = 0 \end{aligned} \quad (40)$$

ここで、

$$P(z|x(t), \theta) \equiv \frac{P(x(t), z|\theta)}{\sum_{z'} P(x(t), z'|\theta)} \quad (41)$$

はデータ $x(t)$ を観測した際の、隠れ変数 z の事後確率 (posterior) と呼ばれる。

非線形方程式である (40) 式をモデルパラメータ θ について解くことは、例えば勾配法を用いれば可能であるが、ここでは (40) 式の形に注目して、以下のような繰り返しアルゴリズムを考える。

1. E (Expectation) ステップ

- (a) 各データ $x(t)$ に対して、現在のパラメータの推定値 $\bar{\theta}$ を用いて隠れ変数 z の事後確率 $P(z|x(t), \bar{\theta})$ を (41) 式により計算する。
- (b) 隠れ変数を含む (完全) データセット $\{(x(t), z(t)) | t = 1, \dots, T\}$ に対する対数尤度 $\sum_{t=1}^T \log P(x(t), z(t)|\theta)$ の、隠れ変数の予測事後確率についての期待値

$$Q(\theta|\bar{\theta}) = \sum_{t=1}^T \sum_{z(t)} P(z(t)|x(t), \bar{\theta}) \log P(x(t), z(t)|\theta) \quad (42)$$

を計算する。

2. M (Maximization) ステップ

期待対数尤度 $Q(\theta|\bar{\theta})$ をパラメータ θ について最大化する。すなわち

$$\frac{\partial Q(\theta|\bar{\theta})}{\partial \theta} = \sum_{t=1}^T \sum_z P(z|x(t), \bar{\theta}) \frac{\partial}{\partial \theta} \log P(x(t), z|\theta) = 0 \quad (43)$$

の解を求める。

3. 求められたパラメータを $\bar{\theta}$ としてステップ 1 に戻る。ステップ 1 と 2 の繰り返しが収束すれば終る。

このアルゴリズムを EM アルゴリズムと呼ぶ。(40) 式と (43) 式の違いに注意する。(43) 式は (40) 式中の一方の θ を $\bar{\theta}$ に置き換えたものである。一般に (40) 式が解析的に解けない場合でも (43) 式は解ける場合があり、それが EM アルゴリズムのメリットである。

E ステップと M ステップを交互に繰り返すことにより、対数尤度が増大することが証明できるため、漸近的にパラメータ θ の最尤推定量を求めることができる。仮に EM アルゴリズムが収束したとすると、 $\theta = \bar{\theta}$ となるが、この時、(43) 式と (40) 式は同じになるので、最尤推定量が求められていることはすぐに分かる。なお、ここで得られる推定量は一般的に局所的最尤解である。

最も簡単な場合の EM アルゴリズムを導出するために、 $\nu_i = 1/M, \sigma_i^2 = \sigma^2 (i = 1, \dots, M)$ の場合の混合正規分布を考える。入力 x の次元は N とする。ここで σ^2 は固定パラメータであり推定はしないものとする。

$$P(x, i|\theta) = \frac{1}{M} (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{1}{2\sigma^2}|x - \mu_i|^2\right) \quad (44a)$$

$$P(x|\theta) = \sum_{i=1}^M P(x, i|\theta) \quad (44b)$$

パラメータは $\theta \equiv \{\mu_i | i = 1, \dots, M\}$ である。

• E ステップ

$$P(i|x(t), \bar{\theta}) = \frac{P(x(t), i|\bar{\theta})}{\sum_{j=1}^M P(x(t), j|\bar{\theta})} = \frac{\exp\left(-\frac{1}{2\sigma^2}|x(t) - \bar{\mu}_i|^2\right)}{\sum_{j=1}^M \exp\left(-\frac{1}{2\sigma^2}|x(t) - \bar{\mu}_j|^2\right)} \quad (45)$$

• M ステップ

$$\log P(x, i|\theta) = -\frac{1}{2\sigma^2}|x - \mu_i|^2 - \log M - \frac{N}{2} \log(2\pi\sigma^2) \quad (46)$$

であるので、

$$\begin{aligned} Q(\theta|\bar{\theta}) &= \sum_{t=1}^T \sum_{i=1}^M P(i|x(t), \bar{\theta}) \log P(x(t), i|\theta) \\ &= -\frac{1}{2\sigma^2} \sum_{t=1}^T \sum_{i=1}^M P(i|x(t), \bar{\theta}) |x - \mu_i|^2 - T \left(\log M + \frac{N}{2} \log(2\pi\sigma^2) \right) \end{aligned} \quad (47)$$

$$= -\frac{T}{2\sigma^2} \sum_{i=1}^M \left(\langle |x|^2 \rangle_i - 2\langle x \rangle_i \mu_i + \langle 1 \rangle_i |\mu_i|^2 \right) - T \left(\log M + \frac{N}{2} \log(2\pi\sigma^2) \right) \quad (48)$$

ここで、

$$\langle 1 \rangle_i \equiv \frac{1}{T} \sum_{t=1}^T P(i|x(t), \bar{\theta}) \quad (49a)$$

$$\langle x \rangle_i \equiv \frac{1}{T} \sum_{t=1}^T x(t) P(i|x(t), \bar{\theta}) \quad (49b)$$

$$\langle |x|^2 \rangle_i \equiv \frac{1}{T} \sum_{t=1}^T |x(t)|^2 P(i|x(t), \bar{\theta}) \quad (49c)$$

を用いた。Q をパラメータ θ (この場合は μ_i) について最大化することは、

$$\frac{\partial Q}{\partial \mu_i} = -\frac{T}{2\sigma^2} (2\langle 1 \rangle_i \mu_i - 2\langle x \rangle_i) = 0 \quad (50)$$

を解くことによって得られる。すなわち

$$\mu_i = \langle x \rangle_i / \langle 1 \rangle_i \quad (51)$$

となる。

結局、E ステップは (45) 式、M ステップは (51) 式で与えられる。この結果についての解釈を与える。(45) 式は Gaussian soft-max と呼ばれる関数である。理解の都合上、分散 σ^2 が小さいとする。(45) 式は、各データ $x(t)$ について、現在の中心位置 $\bar{\mu}_i$ が $x(t)$ にユークリッド距離の意味で最も近いものを選び、そのクラスタへの所属確率をほぼ 1 としている。すなわち現在の中心位置を用いた最小距離規範 (Nearest neighbor 法; NN 法) によるクラスタリングを行なっている。(51) 式では、NN 法によってクラスタリングを行なった後の、各クラスタ内のデータの平均値によってクラスタ中心である μ_i を変更している。(51) 式の分子は各クラスタの構成要素についてデータの和、分子は要素数に対応している。すなわちこの EM アルゴリズムはクラスタリングでしばしば用いられる K 平均法そのものである。違う言い方をすれば、K 平均法は EM アルゴリズムを極めて簡素化したものである。

(演習 3-1) 上記の最も簡単な場合の EM アルゴリズムを用いて実際の 2 次元散布データセット $\{x(t) | t = 1, \dots, T\}$ のクラスタリングを行なう。クラスタの数 M を色々変えて実験をしてみよ。

(演習 3-2) (45) 式において、分散 σ^2 を可変パラメータとした時の EM アルゴリズムを導出せよ。

5 カルマンフィルター

3 章で述べた隠れ変数 (内部状態変数) を含むダイナミクスモデル化の手法の一つとしてカルマンフィルターについて述べる。カルマンフィルタは観測時系列からシステムの内部状態変数を推定する手法として広く用いられている。

対象とするシステムの状態変数が $z(t) \equiv (z_1(t), \dots, z_L(t))$ であり、システムのダイナミクスが以下のよう確率的線形方程式で表わされるものとする。

$$z(t+1) = Az(t) + u(t) + \xi(t) \quad (52)$$

A は $L \times L$ 行列で、 $u(t)$ は時刻 t における外部入力である。 $\xi(t)$ は白色ガウスノイズであり、システムノイズと呼ばれる。すなわち $\xi(t)$ は平均 0、共分散行列 U をもつ正規分布にしたがい、 $t \neq s$ の時、 $\xi(t)$ と $\xi(s)$ は無相関である。以上より、

$$P(z(t+1)|z(t)) \propto \exp \left[-\frac{1}{2} (z(t+1) - Az(t) - u(t))' U^{-1} (z(t+1) - Az(t) - u(t)) \right] \quad (53)$$

が成り立つ。なお本章では正規分布の正規化係数は重要な役割を果たさないで、簡単のために省略する。条件付き確率 (53) は、時刻 t での状態変数が $z(t)$ であった時に、次の時刻 $t+1$ での状態変数が $z(t+1)$ となる状態遷移確率を表わしている。このように、次の時刻の状態がシステムの履歴によらずに現在の状態のみで決まるような確率モデルを、マルコフ過程と呼ぶ。

観測変数 $x(t) \equiv (x_1(t), \dots, x_N(t))$ は状態変数 $z(t)$ から

$$x(t) = Cz(t) + w(t) \quad (54)$$

という関係式で変換されているとする。\$C\$ は \$N \times L\$ の観測行列である。\$w(t)\$ は平均 0、共分散行列 \$V\$ をもつ白色ガウスノイズであり、観測ノイズと呼ばれる。これから

$$P(x(t)|z(t)) \propto \exp \left[-\frac{1}{2}(x(t) - Cz(t))'V^{-1}(x(t) - Cz(t)) \right] \quad (55)$$

となる。

以後、行列 \$A, C, V, U\$ が時刻によらずかつ既知であり、外部入力がない (\$u(t) = 0\$) 場合に、観測時系列 \$\{x(t)|t = 1, \dots\}\$ から内部状態変数 \$z(t)\$ の推定を行なう問題を考える。時刻 \$t\$ までの観測時系列 \$X\{t\} = \{x(s)|s = 1, \dots, t\}\$ をもとに時刻 \$t\$ での状態変数 \$z(t)\$ が、平均 \$\hat{z}(t)\$、共分散行列 \$Q(t)\$ を持つ正規分布にしたがうことがわかっているものとする。すなわち、

$$P(z(t)|X\{t\}) \propto \exp \left[-\frac{1}{2}(z(t) - \hat{z}(t))'Q^{-1}(t)(z(t) - \hat{z}(t)) \right] \quad (56)$$

である。これをもとに、次の時刻 \$t+1\$ での状態変数の分布を推定することができる。

$$P(z(t+1)|X\{t\}) = \int dz(t) P(z(t+1)|z(t)) P(z(t)|X\{t\}) \quad (57)$$

(53) 式と (56) 式から、\$P(z(t+1)|z(t))P(z(t)|X\{t\})\$ は次式のようになる。

$$P(z(t+1)|z(t))P(z(t)|X\{t\}) \propto \exp \left[-\frac{1}{2}E_S(z(t+1), z(t)) \right] \quad (58a)$$

$$E_S(z(t+1), z(t)) = (z(t+1) - Az(t))'U^{-1}(z(t+1) - Az(t)) + (z(t) - \hat{z}(t))'Q^{-1}(t)(z(t) - \hat{z}(t)) \quad (58b)$$

また \$E_S\$ は \$z(t+1)\$ と \$z(t)\$ に関して 2 次形式になっており、以下のように計算することができる。

$$\begin{aligned} E_S(y, z) &= (y - Az)'U^{-1}(y - Az) + (z - \hat{z})Q^{-1}(z - \hat{z}) \\ &= (z - \bar{z})R^{-1}(z - \bar{z}) + (y - A\hat{z})\tilde{Q}^{-1}(y - A\hat{z}) \end{aligned} \quad (59)$$

ここで、

$$\bar{z} = \hat{z} + RA'U^{-1}(y - A\hat{z}) \quad (60a)$$

$$R = (Q^{-1} + A'U^{-1}A)^{-1} \quad (60b)$$

$$\tilde{Q} = U + AQA' \quad (60c)$$

である。(58b)(59) 式より (57) 式中の \$z(t)\$ に関する積分はガウス積分になる。積分を行なうと、(59) 式右辺第 1 項は定数になるので、右辺第 2 項のみが残る。すなわち次式が成り立つ。

$$P(z(t+1)|X\{t\}) \propto \exp \left[-\frac{1}{2}(z(t+1) - A\hat{z}(t))'\tilde{Q}^{-1}(t+1)(z(t+1) - A\hat{z}(t)) \right] \quad (61a)$$

$$\tilde{Q}(t+1) = U + AQA' \quad (61b)$$

\$\tilde{Q}(t+1)\$ は、時刻 \$t\$ における状態分布の知識 (56) とシステムダイナミクスの知識 (53) をもとに、次時刻 \$t+1\$ の状態分布推定を行なった時の状態分布の共分散行列である。(61b) 式からこの共分散が、システムのダイナミクスで定まる値 \$AQA'\$ よりもシステムノイズ分の \$U\$ だけ増大していることがわかる。一方 (61a) 式より、状態変数の期待値は次式で与えられる。

$$\hat{z}(t+1) = A\hat{z}(t) \quad (62)$$

以上は、時刻 t までの観測時系列 $X\{t\}$ をもとに時刻 $t+1$ での状態分布の推定を行なう問題を考えてきた。このように新たな観測を行わずに、ダイナミクスのみを用いて状態分布推定を行なった場合、ノイズのために毎時刻分散が増大してゆき、推定精度が落ちてゆく。また (61) 式より、状態変数の初期分布が正規分布で与えられる場合、以後の状態分布はずっと正規分布になることが分かる。

次に時刻 $t+1$ で新たな観測データ $x(t+1)$ が得られた時に、状態分布の推定がどのように変更されるかを調べる。ベイズの定理より次式が成り立つ。

$$P(z(t+1)|X\{t+1\}) = P(z(t+1)|x(t+1), X\{t\}) = \frac{P(z(t+1), x(t+1)|X\{t\})}{P(x(t+1)|X\{t\})} \quad (63)$$

ここで $P(x(t+1)|X\{t+1\})$ は未知だが、 $z(t+1)$ に依存しないので、 $z(t+1)$ の分布を求める際には単なる定数として扱える。これから

$$P(z(t+1)|X\{t+1\}) \propto P(z(t+1), x(t+1)|X\{t\}) = P(x(t+1)|z(t+1))P(z(t+1)|X\{t\}) \quad (64)$$

が成り立つ。(64) 式の右辺は (55)(61)(62) 式から以下で与えられる。

$$P(x(t+1)|z(t+1))P(z(t+1)|X\{t\}) \propto \exp \left[-\frac{1}{2} E_O(x(t+1), z(t+1)) \right] \quad (65a)$$

$$E_O(x(t+1), z(t+1)) = (x(t+1) - Cz(t+1))'V^{-1}(x(t+1) - Cz(t+1)) \\ + (z(t+1) - \hat{z}(t+1))'\tilde{Q}^{-1}(t+1)(z(t+1) - \hat{z}(t+1)) \quad (65b)$$

E_O に対して、(59) 式の E_S と同様の計算をすることができる。

$$E_O(x, z) = (x - Cz)'V^{-1}(x - Cz) + (z - \hat{z})'\tilde{Q}^{-1}(z - \hat{z}) \\ = (z - \bar{z})'\hat{Q}^{-1}(z - \bar{z}) + (x - C\bar{z})'W^{-1}(x - C\bar{z}) \quad (66)$$

ここで、

$$\bar{z} = \hat{z} + \hat{Q}C'V^{-1}(x - C\hat{z}) \quad (67a)$$

$$\hat{Q} = (\tilde{Q}^{-1} + C'V^{-1}C)^{-1} \quad (67b)$$

$$W = V + C\hat{Q}C' \quad (67c)$$

である。(66) 式の右辺第 2 項は z に依存しないので定数として扱える。結局次式が成り立つ。

$$P(z(t+1)|X\{t+1\}) \propto \exp \left[-\frac{1}{2} (z(t+1) - \hat{z}(t+1))'Q^{-1}(t+1)(z(t+1) - \hat{z}(t+1)) \right] \quad (68)$$

ここで \bar{z} を \hat{z} に書き直した。 $x(t+1)$ を観測した後での $z(t+1)$ の期待値 $\hat{z}(t+1)$ と共分散 $Q(t+1)$ は (67) 式を用いて下のように与えられる。

$$\hat{z}(t+1) = \hat{z}(t+1) + K(t+1)(x(t+1) - C\hat{z}(t+1)) \quad (69a)$$

$$Q(t+1) = (\tilde{Q}^{-1}(t+1) + C'V^{-1}C)^{-1} \quad (69b)$$

$$K(t+1) = Q(t+1)C'V^{-1} \quad (69c)$$

(61b)(62)(69) 式がカルマンフィルタである。以下で式の意味を簡単に説明する。時刻 t での状態の期待値 $\hat{z}(t)$ がわかっている時、システムダイナミクスを用いた予測により $\hat{z}(t+1) = A\hat{z}(t)$ が得られる ((62)

式)。この予測は観測データ $x(t+1)$ を用いて修正される。すなわち、観測データと内部状態の予測値 $\hat{z}(t+1)$ から予測される観測値との差 $(x(t+1) - C\hat{z}(t+1))$ を用いて $\hat{z}(t+1)$ が修正される ((69a) 式)。この修正のための比例係数行列 $K(t+1)$ はカルマンゲインと呼ばれる。

状態分布の共分散の大きさがどのように変化するかを見るために、特に 1 次元の場合を考える。すなわち、システムノイズの分散 $U = \sigma_S^2$ 、観測ノイズの分散 $V = \sigma_O^2$ 、 $z(t)$ の分散を $Q(t) = \sigma(t)^2$ とする。(61b) 式と (69b) 式より

$$\sigma^2(t+1) = \left(\frac{1}{\sigma_S^2 + A^2\sigma^2(t)} + \frac{C^2}{\sigma_O^2} \right)^{-1} \quad (70)$$

となる。この式から観測を行なった後での分散 $\sigma^2(t+1)$ は観測を行う前の分散 $\tilde{\sigma}^2(t+1) = A^2\sigma^2(t) + \sigma_S^2$ よりも小さくなっていることがわかる。また観測ノイズの分散 σ_O^2 が非常に小さい時 $\sigma^2(t+1) \approx \sigma_O^2/C^2$ となり状態の分散も小さくなり推定精度が上がる事がわかる。

これまで A, C, V, U は既知だとしてきたが、未知の場合でも EM アルゴリズムを用いて状態推定とパラメータ推定を同時に行なうことができる。

(演習 4) システムノイズと観測ノイズがある状況での単振子の状態推定問題に対してカルマンフィルタを適用する。振子の鉛直方向からの角度を θ とすると、 θ があまり大きくない場合の振子の運動方程式は

$$\frac{d^2\theta}{dt^2} = -\omega^2\theta, \quad \omega^2 = \frac{g}{r}$$

で与えられる。ここで g は重力加速度、 r は振子の長さである。角速度を $v \equiv d\theta/dt$ で定義すると運動方程式は、

$$\frac{d\theta}{dt} = v, \quad \frac{dv}{dt} = -\omega^2\theta$$

となる。時間間隔 Δt でサンプリングして、オイラー差分により離散時間方程式に変換すると、

$$\begin{aligned} \theta(n+1) &= \theta(n) + \Delta t \cdot v(n) \\ v(n+1) &= v(n) - \Delta t \cdot \omega^2\theta(n) \end{aligned}$$

となる。分散 σ_S^2 をもつシステムノイズと分散 σ_O^2 をもつ観測ノイズとが付加されたシステムから、(ノイズの付加された) 角度の時系列 $\{\theta(n)|n = 1, \dots\}$ が観測できる状況で、状態変数である $\theta(n)$ と $v(n)$ を推定せよ。

2 日目 ダイナミクスの予測 小脳と内部モデル

コーディネータ：片山正純，北澤茂

1．講義のねらいと概要

2 日目は「小脳」，キーワードは「内部モデル」である．

講義 1, 2, 3 は小脳内部モデル理論の集大成であり，参加者は科学史の展開に現在進行形で立ち会う興奮を感じるだろう．講義 1 では運動制御における内部モデル（順モデルと逆モデル）の定義と役割，さらに，内部モデルを学習により獲得する学習制御モデルについて工学的な知識が与えられる．講義 2 では，小脳に内部モデルが学習によって獲得されるという「小脳内部モデル理論」が精密な実験データによって裏付けられていく．まさに科学の王道ともいべき展開が示される．さらには，多重順逆内部モデルなどの最新の理論的發展にまで話は及ぶ．講義 3 では，実験心理学の手法と脳機能イメージングの方法を組み合わせ，「人間の」小脳に内部モデルが獲得される様子を計測した最新の研究が紹介される．いよいよ「人間の脳」がハードサイエンスの対象となったのである．

講義 4, 5 は運動の中でも到達運動に焦点をあてる．講義 4 では，到達運動に潜む美しい性質が，小脳で得られたボトムアップのデータでどこまで説明できるか，ボトムアップの実験屋からトップダウンの理論家への発信を試みる．講義 5 では到達運動を例にとって，独創的な問題意識に基づく話題が提供される．射程は小脳の教師つき学習と基底核の強化学習の関係に及び，3 日目への橋渡しともなる．活発な議論が今から楽しみである．

2．必読文献

1. Kawato, M. (1999) Internal models for motor control and trajectory planning, *Current Opinion in Neurobiology*, 9: 718-727.
2. Imamizu, H. et al. (2000) *Nature*, 403: 192-195.
3. Kitazawa, et al. (1998) *Nature*, 392: 494-497
4. Houk JC, Wise SP (1995) Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. *Cerebral Cortex*, 5: 95-110.

内部モデル学習制御と腕の柔らかさ調節の統合に関して

片山正純(豊橋技術科学大学)

1. はじめに

腕の随意運動 (Voluntary Movement) に関する運動制御機構の研究は古くから行われているが、フィードバック制御機構 (Feedback Control Mechanism) に基づいて説明されてきた。しかし、20年程前からフィードフォワード制御機構 (Feedforward Control Mechanism) の存在が生理学や行動学の分野で実験的に確かめられた (例えば, Polit ら, 1979; Bizzi ら, 1984)。一方, 生体の運動制御系には, 神経伝達, 神経情報処理, 筋や固有受容器 (筋紡錘, ゴルジの腱器官など) などにより生じる時間遅れが存在するために, フィードバック制御のループ時間 (1回の制御ループに必要な時間) が大きくなってしまふ。例えば, 単シナプス性の反射 (伸張反射など) でさえ約30ミリ秒必要であり, さらにトランスコーティカルループ (大脳皮質を介して閉ループ制御系) ではそのループ時間は50ミリ秒以上必要となり, 視覚情報によるフィードバック系では少なくとも150ミリ秒程度以上必要となる。このため, 大きすぎるフィードバックゲインは運動を振動的にさせたり不安定にさせるため, 単純なフィードバック制御だけで腕を安定に精度良く制御することが困難である。従って, 腕の運動制御機構において, フィードバック制御機構だけでなくフィードフォワード制御機構が重要な役割を果たしている。

以上の観点から, フィードフォワード制御機構に関する研究が1980年頃から盛んに行なわれるようになり, これらの研究は2種類のアプローチに大別することができる。1つは, 神経・筋骨格系のバネのような特性 (粘弾性特性) を直接利用することにより逆動力学 (Inverse Dynamics) の計算を行わないでフィードフォワード制御機構を説明するアプローチである (内部モデルを用いないアプローチ)。この代表的な制御機構として仮想軌道制御仮説 (Virtual Trajectory Control Hypothesis) がある (Bizzi ら, 1984; Hogan, 1984; Flash ら, 1987)。もう1つは, 筋骨格系の内部モデル (逆動力学モデル) を学習することによってフィードフォワード制御機構を説明するアプローチである (ここでは内部モデル学習制御と呼ぶことにする)。この代表的な学習制御機構としてフィードバック誤差学習法により逆動力学モデルを獲得する階層神経回路モデルがある (Kawato ら, 1987)。最近, 仮想軌道制御仮説に基づいた制御では, 仮想軌道が複雑になることが示され (Katayama and Kawato, 1993; Gomi and Kawato, 1997), 仮想軌道制御仮説においても逆動力学の問題を解いて仮想軌道を計画する必要があることが明らかになった。従って, どちらの制御仮説においても内部モデル (逆動力学モデルなど) を用いたフィードフォワード制御機構が重要な役割を果たしていると言える。この観点から, 内部モデルの役割, 内部モデルを用いた制御, 内部モデルの学習について説明する。

さらに、内部モデル学習制御と腕の柔らかさ（筋の粘弾性）調節の統合の可能性について議論する。従来より提案されてきた内部モデル学習制御モデルでは、ヒトの運動制御系の性質（筋の可変粘弾性など）がほとんど考慮されてこなかったため、運動制御系の性質を考慮した内部モデル学習制御メカニズムを議論する必要があると考えている。この観点から、運動系の特徴の1つである筋の可変粘弾性に着目し、より効果的に内部モデルを学習・制御するための順逆モデルを用いた学習制御モデルについて議論し、計算論的な利点についても考察したい。

2. 3つの計算問題

目標まで手先を伸ばす運動（到達運動）をフィードフォワードで実現するためには、腕の運動軌道を計画する「軌道生成」の問題、作業座標系（または視覚座標系）から関節角や筋長など身体座標系に変換する「座標変換」の問題、計画した軌道を正確に実行するための運動指令（運動ニューロンや筋の活性化レベルを決める指令値）を求める「制御」の問題が少なくとも解かれなければならない（図1参照）。ロボティクスでは、座標変換の問題を逆運動学（Inverse Kinematics）、制御の問題を逆力学（Inverse Dynamics）と呼ぶ。これらの問題は解が一意に定まらないという意味で不良設定問題である（図2参照）。つまり、軌道生成では腕の軌道が無数に存在し、座標変換ではある手先位置を実現する腕の姿勢は無数に存在する。さらに、関節トルクは屈筋と伸筋の筋張力の差で生成される。このため、制御では、運動に必要な関節トルクを生成するための運動指令の組み合わせは無数に存在する。以上のような不良設定性の存在は、運動軌道、腕の姿勢、腕の柔らかさなどが運動の種類や目的に応じて調節可能であることを意味している。つまり、人は何らかの拘束条件や最適化原理に基づいて上記問題を解決し、運動や対象物操作を巧みにやっていると考えられる。ここでは内部モデルを用いることにより「制御の問題」を解決するための基本理論に焦点をあてて解説する。

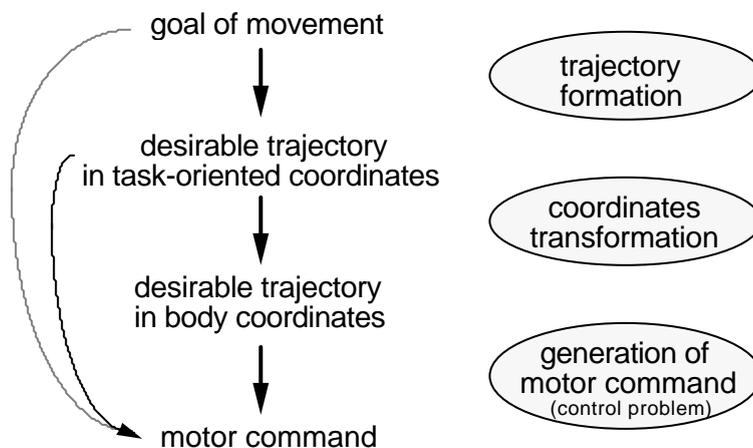


図1 3つの計算問題

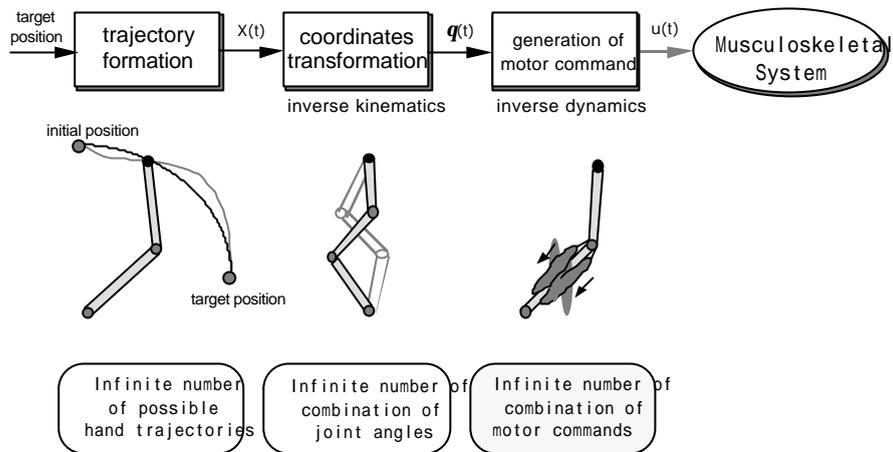


図2 3つの計算問題における冗長性

3 . 内部モデル制御

Ito(1970)は小脳に内部モデルを形成していることを30年も前から指摘しており，Marr (1969)とAlbus (1971)は同時期に小脳パーセプトロン仮説を提案している．その後1980年ころから内部モデルを用いたフィードフォワード制御メカニズムのモデル化が盛んに行われるようになった．さらに，Kawatoら(例えば，1987)は，ItoやAllen&Tsukahara (1974)の研究を発展させて，大脳小脳連関，赤核，小脳の神経回路モデルとして，順モデルと逆モデルを用いた制御メカニズムを提案した(図3)．このように，学習制御系では，内部モデルとして順モデルや逆モデルの存在が指摘されており，重要な役割を果たしていると考えられている．順モデルの主な役割は，運動指令の遠心性コピーを用いて，1) 次の状態や実現軌道を予測する，2) 内部フィードバック制御を構成する，などが考えられ，さらに後述の誤差変換(図6(b))が考えられる(例えば，Wolpert (1997)やKawato (1999))．一方，逆モデルの主な役割は，1) 制御対象のダイナミクスを補償したフィードフォワード制御を構成する，などである(例えば，川人(1996)やKawato (1999))．

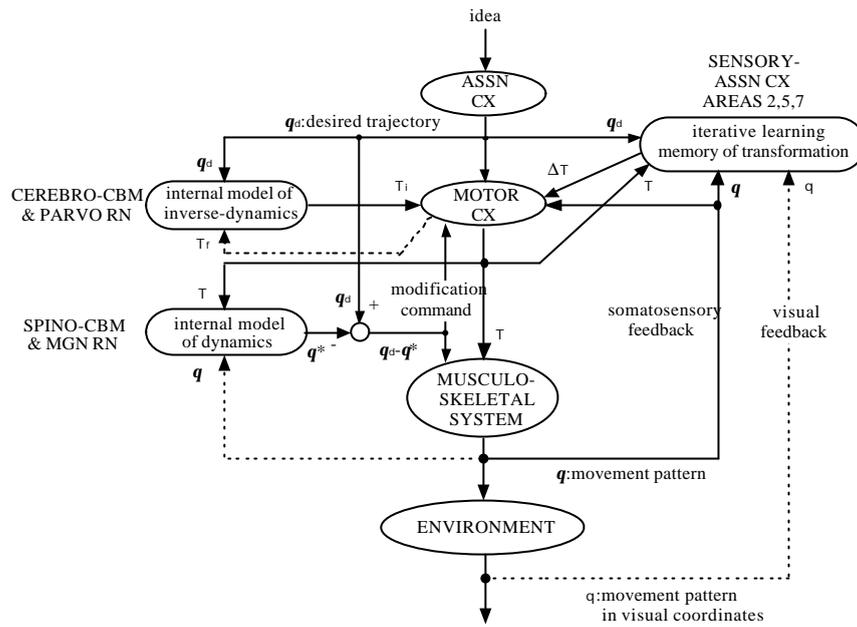
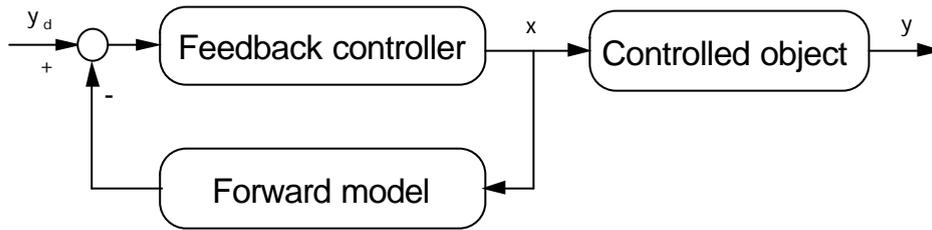
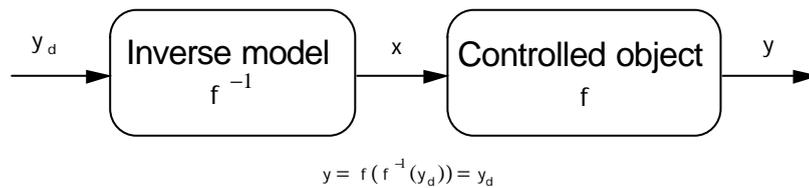


図3 随意運動制御スキーム



(a) control scheme using a forward model

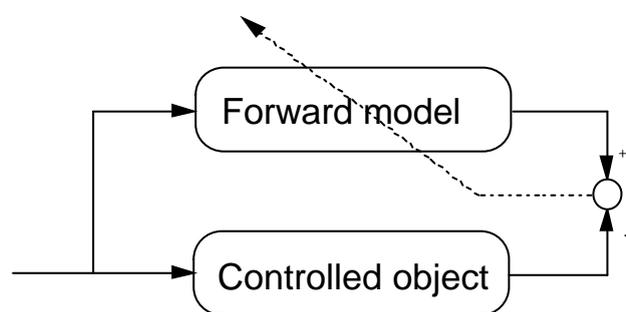


(b) feedforward control scheme using an inverse model

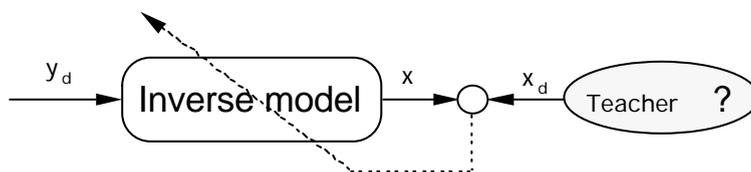
図4 順モデルと逆モデルを用いた制御

そこで、まず順モデルを用いた内部フィードバック制御について説明する(図4(a)参照)。順モデル(正確には順動力学モデル)は制御対象(Controlled Object)と同じ入出力関係をもつモデルである。入力は運動指令(関節トルクや筋への指令など)であり、出力は実現軌道(関節角度や筋長など)である。前述のようにヒトの運動制御系には大きな時間遅れが存在する。この

ため、フィードバック制御だけでは目標軌道を実現できないが、図4(a)のように順モデルを用いた内部フィードバックループを構成することにより、目標軌道を達成できるようになる。この制御系では感覚フィードバックを用いていないため、大きな時間遅れ無しに制御できるためである。しかし、順モデルを用いた内部フィードバックループでも10〜30ミリ秒程度の遅れが存在し、さらに順モデルの出力誤差が積分されていく構造になっているため順モデルだけを用いた内部フィードバックループだけでは正確に目標軌道を実現することは困難である。次に、逆モデルを用いた制御について説明する(図4(b)参照)。逆モデル(正確には逆力学モデル)は制御対象とは逆の入出力関係をもつモデルである。図4(b)に示すように、制御対象を関数 f とすると逆モデルは逆関数 f^{-1} となる。つまり、これらの合成写像を考えると、 $x=f^{-1}(y)$ 、 $y=f(x)$ より、 $y=f(f^{-1}(y_d))=y_d$ となり、正確に目標軌道 y_d が正確に実現されることになる。つまり、逆モデルに目標軌道を入力して出てきた出力を制御対象に与えることにより正確に目標軌道が実現できるのである。従って、逆モデルは理想的なフィードフォワード制御器である。また、ヒトの筋骨格系は非線形性が強いいため、特に逆モデルを用いた制御は有効である。最後に、対象物操作においても本節において説明した制御スキームは有効である。つまり、制御対象(腕など)の順・逆モデルだけでなく、操作対象の順・逆モデルを用いることにより、上述の制御は可能となる。



(a) forward model learning



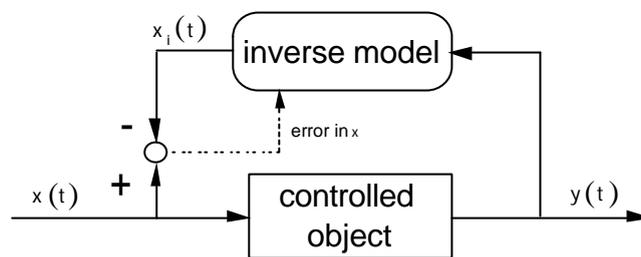
(b) inverse model learning

図5 順モデルの学習と逆モデルの学習の困難さ

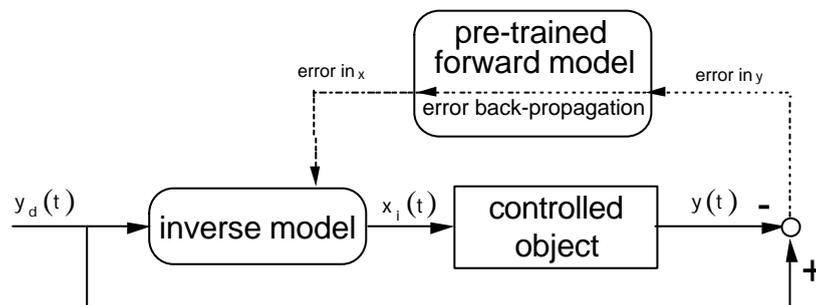
4 . 内部モデル学習

内部モデル(順モデルと逆モデル)は生得的に脳内に獲得されているとは考えにくい。そこで、

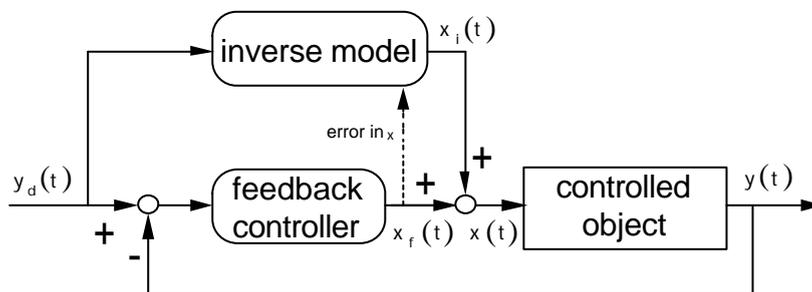
本節では、順モデルと逆モデルの学習について説明する。順モデルは、図5(a)に示すように、制御対象と順モデルに同じ入力を与え、そのそれぞれの出力の誤差を用いて学習することができる。つまり、この誤差を小さくするように順モデルのパラメータを更新すればよい(誤差逆伝搬法など)。一方、逆モデルの学習について説明する。図5(b)に示すように、逆モデルの教師信号(正解値)が既知の場合には逆モデルの出力と教師信号との誤差を用いて学習することができる。しかし、教師信号が既知であれば逆モデルを学習する必要はないし、一般的には逆モデルの出力(筋の活動度を決める運動指令)の教師信号は存在しない。観測できるのは実現した軌道だけである。そこで、この困難さを解決した3種類の学習スキームが提案されている。



(a) direct inverse modeling



(b) forward and inverse modeling



(c) feedback error learning

図6 逆モデルの3つの学習スキーム

ここでは、これらの3つの学習スキームについて簡単に説明する(詳しくは川人(1996)を参

照してください)。まず、最もシンプルな学習スキーム(図6(a))について説明する。この学習スキームでは、制御対象の出力を逆モデルに入力し、制御対象への入力と逆モデルの出力との誤差を用いて学習することができる。この学習スキームは Jordan によって直接逆モデリング(direct inverse modelling)と呼ばれている。この手法はもっともシンプルであるため簡便である。しかし、いくつかの欠点がある。例えば、(1)目標指向性が無い、(2)制御対象に冗長性がある場合には使用できない、(3)学習と制御で逆モデルの接続を変更する必要がある、などである。次に、Jordanら(1992)によって提案された順逆モデリング(forward and inverse modelling)について説明する(図6(b))。この学習スキームでは、順モデルを用いることにより、軌道での誤差を逆伝搬させることにより運動指令での誤差に変換する。この変換された運動指令の誤差を用いて逆モデルを学習することができる。この学習スキームでは、直接逆モデリングの欠点(1-3)を解決しているが、予め順モデルを学習しておく必要がある。最後に、Kawatoら(1987)によって提案されたフィードバック誤差学習(feedback-error learning)について説明する(図6(c))。この学習スキームは、図3における大脳小脳連関、運動野、小脳外側部、小細胞性赤核、トランスコーティカルループの部分モデル化したものである。この学習スキームでは、フィードバックコントローラの出力を逆モデルの出力誤差として用いることにより、逆モデルを学習する。つまり、フィードバックコントローラの出力が小さくなるように逆モデルを学習するのである。このため、学習前ではフィードバックコントローラ主体で制御されるが、学習するにつれて逆モデルの出力を用いたフィードフォワード制御に移行する。理想的には学習後に逆モデルだけを用いたフィードフォワード制御が可能となる。この学習スキームでも直接逆モデリングの欠点(1-3)を解決している。産業用マニピュレータやゴム製の人工筋(空気圧駆動)をもつマニピュレータの学習制御に成功している。特に、後者の場合には200-300ミリ秒程度の時間遅れが存在するにも関わらず精度の良い制御が可能となっている(片山ら(1990))。

5. 内部モデル学習制御と腕の柔らかさ調節の統合

上述のように、内部モデルを用いたフィードフォワード制御系に関して様々な研究がなされてきた。しかし、これらの学習制御モデルでは、運動制御系の調節メカニズムが考慮されていない。つまり、2章で説明したようにいろいろなレベルで冗長性が存在しているため、運動軌道、腕の姿勢、腕の柔らかさなどを運動の目的や学習の達成レベルに応じて調節できることを意味している。ヒトが経験したことのない運動やスポーツを学び始めたときでも、何とか合目的に運動するであろう。つまり、運動軌道、腕の姿勢、腕の柔らかさなどを調節することにより学習や制御をより簡単にすることにより、学習初期でもある程度目的が達成できるように運動することができるのである。これは運動学習の1つの戦略である。このように、運動制御系における調節メカ

ニズムと内部モデル制御を統合することにより、よりもっともらしい学習制御モデルが構築できる可能性がある。

そこで、本節では、腕の柔らかさ（粘弾性）調節に着目し、より効果的に内部モデルを学習・制御するための順逆モデルを用いた学習制御モデルについて考えてみたい。まず、非常にシンプルな粘弾性の調節メカニズムについて説明する。基本的なアイデアは、運動誤差に応じて、粘弾性を変化させることである。具体的には、体性感覚フィードバック情報と目標軌道との誤差（運動誤差）と順モデルにより予測した軌道と目標軌道との誤差（予測誤差）の合計に応じて粘弾性（屈筋と伸筋の同時活性化レベルなど）を変化させることである。フィードバックコントローラのゲインを大きくすると大きな時間遅れのため腕が振動したり不安定になる。しかし、腕の粘弾性を上げる（腕を硬くする）ことはフィードバックコントローラのゲインを大きくしたときと同様の効果があるため、内部モデルが学習できていないときでも粘弾性を大きくすることにより、ある程度の精度で運動することができるようになる（Katayama ら（1998））。この基本的アイデアに基づいて構成した学習制御モデルを図7に示す。この学習制御モデルでは、2章と3章で説明した順モデルを用いた内部フィードバック制御とフィードバック誤差学習を組み合わせている。また、運動誤差 e_{fb} に基づいて粘弾性を調節する Feedback Adjustor と予測誤差 e_{pr} に基づいて粘弾性を調節する Predictive Adjustor の合計により、粘弾性レベル C （または同時活性化レベル）を決定する。 g は単調増加関数である。

$$C(t) = g(e_{fb}(t) + e_{pr}(t))$$

フィードバック誤差学習では学習を繰り返すにつれて運動誤差が徐々に減少するため、この調節則に従って学習を行うことにより、粘弾性レベルも徐々に小さくなる。学習前には粘弾性レベルが大きくなっているために、疲労や消費エネルギーの観点から望ましくないが、学習後にはより小さな粘弾性レベルで正確に運動できるようになる。

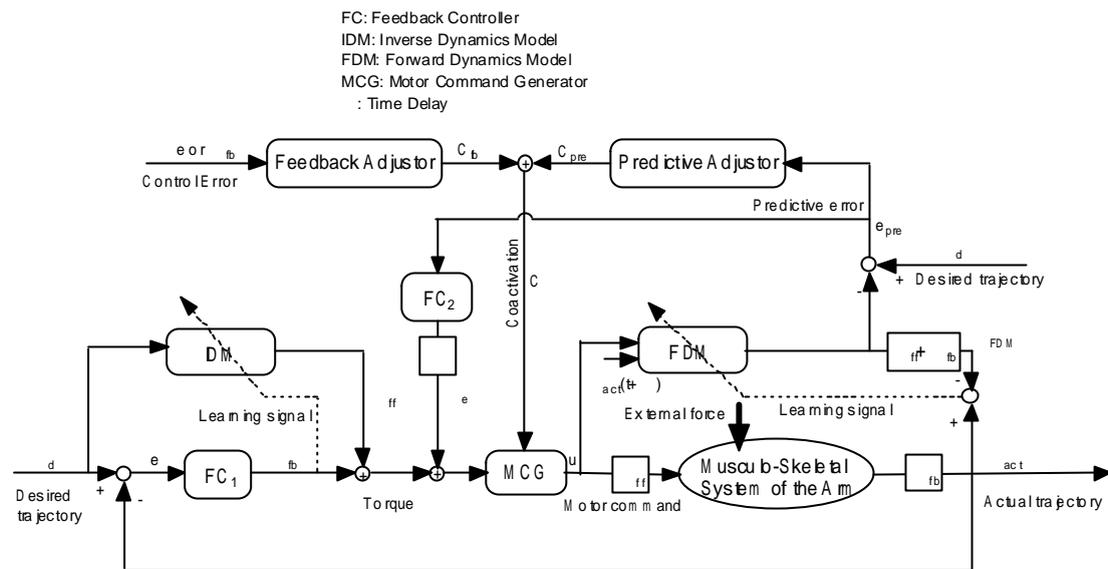


図7順・逆モデルを用いた粘弾性調節による学習制御モデル

この学習制御スキーム（粘弾性の調節法）を用いた利点について説明する．まず，前述のように学習初期でもある程度の精度で運動可能となる．また，粘弾性レベルを大きくすることにより，静的な力学関係に関する逆モデル（逆静力学モデル；片山，川人，1990）だけで制御可能となる．このため，この調節則を用いて学習・制御することにより，より簡単な学習・制御から徐々に難しい学習・制御に移行することになり，効率の良い学習が期待できる．さらに，学習初期に粘弾性レベルを大きくすることにより，正しい解周辺から学習を開始するため，局所解へ落ち込む確率を小さくしたり，学習の収束を早くするなどが期待できる．また，学習過程において，粘弾性（または筋活動）が徐々に小さくなることは実験的に調べられている（例えば，大須ら（1997）やThoroughmanら（1999）など）．また，順逆モデルの必要性についてもBhushanら（1999）によって示されている．さらに，この学習制御モデルは順逆モデルを用いているためWolpert&Kawato（1998）により提案された多重順逆モデルの制御スキームに基づいて構成することも可能である．

6. おわりに

本稿において，内部モデル学習制御に関する基礎理論について川人（1999）の文献に沿って1章から4章で簡単に説明した．詳しくは川人（1999）を熟読してください．また，内部モデル学習制御と腕の柔らかさ（筋の粘弾性）調節の統合の可能性について簡単に説明した．このように，運動制御系の調節メカニズムを考慮した運動学習メカニズムを考えることにより，様々なレベルでの計算論的な利点があるだけでなく，より妥当なモデル化が可能となるのではないかと考えている．すなわち，本稿では腕の柔らかさ調節に焦点を絞って議論したが，運動軌道や腕姿勢，

さらには様々なレベルでの反射系などの調節メカニズムを考慮した運動学習メカニズムを考えてみるのも面白いのではないかと考えている。今後、心理物理実験や計算機シミュレーションにより、より詳細に調べてゆきたいと考えている。

必読文献

Kawato M: Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology* 9, pp.718-727 (1999).

参考文献

Albus, J.S. (1971) A theory of cerebellar function. *Math. Biosci.* 10, 25-61

Allen, J.S., Tsukahara, N. (1974) Cerebrocerebellar communication systems. *Physiological Review* 54, 957-1006.

Bhushan N, Shadmehr R: Computational nature of human adaptive control during learning of reaching movements in force fields. *Biological Cybernetics* 81, pp.39-60 (1999).

Bizzi, E., Accornero, N., Chapple, W., Hogan, N. (1984) Posture control and trajectory formation during arm movement. *The Journal of Neuroscience* 4, 11, 2738-2744.

Flash, T. (1987) The control of hand equilibrium trajectories in multi-joint arm movement. *Biological Cybernetics* 57, 257-274.

Gomi, H. and Kawato, M. (1997) Human arm stiffness and equilibrium-point trajectory during multi-joint movement. *Biological Cybernetics*, 76, 163-171.

Hogan, N. (1984) An organizing principle for a class of voluntary movement, *The Journal of Neuroscience*, 4, 11, 2745-2754.

Ito, M. (1970) Neurophysiological aspects of the cerebellar motor control system. *International Journal of Neurology*, 7, 162-176.

Jordan, M.I., Rumelhart, D.E. (1992) Forward models: Supervised learning with a distal teacher, *Cognitive Science* 16, 307-354.

片山正純, 川人光男(1990) 筋肉 骨格系の運動制御を行う並列階層制御神経回路モデル, 電子情報通信学会論文集(ニューロコンピューティング特

Katayama, M., Kawato, M. (1993) Virtual trajectory and stiffness ellipse during multi-joint arm movement predicted by neural inverse models. *Biological Cybernetics* 69: 353-362.

Katayama M, Inoue S, Kawato M: A strategy of motor learning using adjustable parameters for arm movement. *Proceedings of the 20th Annual International Conference on the IEEE Engineering in Medicine and Biology Society*, pp.2370-2373 (1998). Parahewa S, 片山正純: 予測調節機構を用いた内部モデル学習制御, 電子情報通信学会技術研究報告 NC99, pp.23-30 (2000).

Kawato, M., Furukawa, K., Suzuki, R. (1987) A hierarchical neural-network model for control and learning of voluntary movement, *Biological Cybernetics*, 57, 169-185.

川人光男：脳の計算理論，産業図書（1996）（運動制御に関わる章：4章，5章，6章）

Marr, D. (1969) A theory of cerebellar cortex. *Journal of Physiology*, 202, 437-470.

集号) J73-D-II, 8, 1328-1335.

大須，道免，五味，吉岡，今水，川人：運動学習時における筋活性の変化．電子情報通信学会技術研究報告 NC96-139，pp.201-208（1997）.

Polit, A., Bizzi, E. (1979) Characteristics of the motor programs underlying arm movements in monkeys. *J. Neurophysiol.* 42, 183-194.

Thoroughman KA, Shadmehr R: Electromyographic correlates of learning an internal model of reaching movements. *The Journal of Neuroscience* 19, 19, pp.8573-8588 (1999).

Wolpert DM: Computational approaches to motor control. *Trends in Cognitive Sciences*, 1, 6, pp.209-216 (1997).

Wolpert DM, Kawato M: Control of Arm and Other Body Movements - Multiple paired forward and inverse models for motor control. *Neural Networks* 11, 7, pp.1317-1330 (1998).

小脳に学習で獲得される内部モデル

川人光男

A T R 人間情報通信研究所

科学技術振興事業団、創造科学推進事業、川人学習動態脳プロジェクト

アブストラクト

小脳は、運動制御だけでなく、言語や思考を含むヒト高次認知機能に重要な働きをすることがわかってきた。小脳の機能は原始的な反射から高次機能まで実に多種多様であるが、小脳皮質の回路構造とシナプス可塑性は一様であるから、どの機能においても共通の計算原理が働いていると考えられる。我々の小脳内部モデル理論では、“小脳皮質に登上線維入力伝える誤差信号に誘導され、プルキンエ細胞のシナプス可塑性に基づいて、外界の内部モデルが学習で獲得される”と考える。この理論の予測は、サルを対象にした電気生理学的研究と、ヒトの脳活動の非侵襲計測研究によって支持されている。

はじめに

前回の講演では、ヒト知性に迫る計算論的神経科学のアプローチを概観した。このアプローチでは、これからくわしく述べていく“階層的多重順逆モデル”が中心的な役割を果たした。今回紹介するのは、このモデルの基礎をなす、小脳の内部モデル仮説である。

小脳は運動制御のためにあるというのがこれまでの常識であった。ところが最近の研究は、この考えを根底からくつがえし、ヒト小脳は高次認知機能で活動することがわかった⁽¹⁾。小脳の異なる部位は、脳、脊髄などの多様な神経部位と結びつき、この多様な結合と対応してさまざまな機能とかがかかわっている。自律神経系の調節、反射運動の適応、条件づけ、歩行運動、随意運動の実行と計画、運動の想像、ゲームの思考、暗算、視覚や体性感覚に基づくパターン認識、メンタルローテーション、言語など実に多種多様で、原始的な機能から、ヒト特有の知性にまで及ぶ研究が続々と報告されている。

ところが、小脳皮質の神経回路は、小脳のどの部位でも一様である。これは、大脳皮質の異なる領域が、ニューロン（神経細胞）の種類だけでなく、層構造まで異なるのと対照的だ。小脳研究の魅力は、そのハードウェアの一様性と機能の多様性のコントラストにあるのだ。小脳皮質のシナプス可塑性（刺激の組合せによっておこるシナプス伝達効率の変化）も小脳内の部位によらない。小脳皮質の独特の神経回路と可塑性や他の脳部位との結合様式から考えて、その情報処理には大脳皮質、大脳基底核、海馬の機能とは異なる小脳に特有の何かがあるはずである。単純な運動制御から言語まで、学習によって小脳に獲得される内部モデルが、その特有の何かだといえることを示そう。

小脳皮質の神経回路

小脳は、脊椎動物の脳の一部をさし、ヒトでは大脳皮質の後頭葉の下に半分隠れるように位置する。ヒトでは、大脳に比べて、体積は 1/10 であるが、表面積は半分以上、さらにニューロンの数はかえっ

て多い。小脳皮質は、わずか 5 種類のニューロンから構成されている（プルキンエ細胞、顆粒細胞、ゴルジ細胞、バスケット細胞、星状細胞）。

小脳皮質からの出力をおこなうのは、そのうちプルキンエ細胞に限られ、小脳核へと投射する。そのプルキンエ細胞には、主に 2 種類の線維をとおしたシナプス入力がある（図 1）。プルキンエ細胞 1 個当たり、数十万個のシナプスを作る平行線維と、ただ 1 本であるが強力なシナプスを作る登上線維である。違いはシナプスの数だけではない。平行線維は、プルキンエ細胞に数百パルス / 秒にまでいたる高

発火頻度の単純スパイク（スパイクとは、電気的なパルスのこと）をひきおこす。いっぽう、登上線維は、1 秒間にたかだか 1 個か 2 個しか生じない複雑スパイクをひきおこす（図 1）。

平行線維からのシナプスの伝達効率は、登上線維からの信号に依存して、可塑的に変化する。平行線維入力と登上線維入力

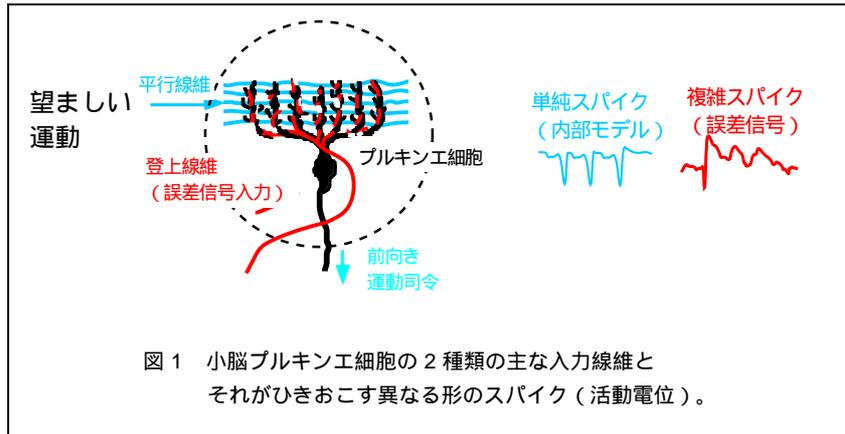


図 1 小脳プルキンエ細胞の 2 種類の主な入力線維とそれがひきおこす異なる形のスパイク（活動電位）。

同時に興奮（発火）するとそのシナプスの伝達効率は減少していくことになる（長期減弱⁽²⁾）。いっぽう、登上線維が発火していないのに、平行線維だけが興奮するとシナプス伝達効率は増加していく（長期増強）。この 2 種類のシナプス可塑性が対になって働くことによって、登上線維入力によるプルキンエ細胞の発火は、興奮しすぎるでもなく、興奮しなさすぎるでもない、ちょうどよい程度に保障されている。これはいいかえれば、登上線維がもし学習のための誤差信号を表わしているなら、理論的によく知られている教師あり学習則がシナプス可塑性のしくみによって、実現されていることになる。つまり、プルキンエ細胞の発火が調節されて誤差を減少させている。

小脳は、苔状線維（顆粒細胞に向かう）の起始細胞からなる層、顆粒細胞からなる層、プルキンエ細胞からなる層の 3 層を持ち、その間に主に前向き（起始細胞（苔状線維）顆粒細胞 プルキンエ細胞）の結合をもつ。この 3 層神経回路の第 2 層（中間層）を構成する顆粒細胞が 100 億個以上も存在することなどに着目した最近の数学的（理論的）な研究によって、この回路でかなり広い範囲の非線形関数が実現できることがわかる。

小脳皮質は、数千個から、数万個のマイクロゾーンと呼ばれる小さな解剖学的単位に別かれ、一つのマイクロゾーンに、同じ性質を持つ登上線維が入力する。小脳皮質の中でも系統発生的に新しい部位に向かう苔状線維の起始細胞は橋核にあり、その主な入力は大脳皮質から来る。新しい小脳部位からの出力は、小脳核の一つ歯状核をとおり、視床を介して、大脳皮質に戻る。したがって、系統発生的に新しい小脳皮質は、大脳と、入出力ともにループ回路を構成している。

大脳皮質と橋核を結ぶ大脳皮質橋線維は 2000 万、苔状線維は 2000 万、顆粒細胞は 100 億、プルキンエ細胞は 2000 万、小脳核細胞は 200 万個ある。マイクロゾーンが約 5000 あるとすると、一つのマイクロゾーンあたり、大脳皮質橋線維が 4000 本、苔状線維が 4000 本、顆粒細胞が 1000 万個程度、

プルキンエ細胞が 4000 個、小脳核細胞が 400 個程度あることになる。出力（小脳核細胞）の数が、入力（大脳皮質橋線維）の数のほぼ 1/10 であるから、小脳では情報を縮約しながら、入力を出力に変換していることになる。

内部モデル

小脳皮質が、登上線維入力を教師信号として、運動学習をおこなうとする仮説（Marr-Albus-Ito 仮説⁽³⁾⁽⁴⁾⁽⁵⁾）は、1970 年前後に提案された。その後、学習によって制御対象の内部モデルが小脳皮質に獲得されるとする計算理論が提案されたのは、1980 年代終わりころである。

ここでいう内部モデルとは、脳の外に存在するある対象の入出力特性をまねることができる脳内の神経回路のことだと考えてほしい。われわれヒトを始めとする動物の脳の中にあるということで、内部という用語を用いる。外界のある物のまねをする、シミュレーションする、エミュレーションするという意味で、モデルという用語が使われている。

ヒトが身体を動かすとき、脳はかなり複雑な計算問題を解く必要があり、内部モデルを必要とする⁽⁶⁾。ヒトの身体や環境の動特性は、生後激しく変化するので、あらかじめ内部モデルを遺伝的に決められない。したがって、内部モデルは、シナプス可塑性にもとづいて学習で獲得しなければならないことになる。内部モデルは、実際にはどのようなものなのだろうか。

内部モデルには、順モデルと逆モデルの 2 種類がある。順モデルとは、モデリングの対象となるダイナミカルシステムと同じ入出力の方向性をもち、同じ入出力変換を実現するものである。例えば、腕や眼球などの運動制御の対象は、筋肉の張力などの運動指令を受け取って運動軌道（腕や眼球の動き）を実現するが、それと同じ入出力特性をもつモデルを身体の順モデルと呼ぶ。順モデルは実際に運動システムに送られる運動指令のコピーを受け取り、実際に実現される軌道の推定をおこなう。

いっぽう、逆モデルとは、ダイナミカルシステムの入出力をひっくり返した、いわば入出力変換の逆関数をシミュレーションするものである。身体の逆モデルは、運動軌道が入力されるとそれを実現するために必要な運動指令を計算する。その入出力特性は制御対象の入出力特性のちょうど逆になっているので、逆モデルと制御対象とを直列につなぐと恒等写像になる。このシステムに目標軌道を入力すると、実現される軌道は目標軌道と同じになる。つまり、逆モデルはフィードバック情報を使わない前向き制御のための理想的な制御器になっている。

過去 15 年ほどの人工的な神経回路モデルの研究で、内部モデルを定義する一見複雑な式が簡単な構造をもつ人工的な神経回路モデルで実現されることがわかってきた。とくに、逆モデルは簡単な 3 層の前向き神経回路で実現できる。この人工的な神経回路は、実際の小脳の神経回路に比べればずっと簡単な構造をしているから、小脳に、われわれが考えているような逆モデルや順モデルがあっても何の不思議もない。内部モデルの学習は、モデリングの対象が外界にあり、例えば順モデルの学習では誤差信号が内部モデルによる推定と実際の感覚フィードバックの差として簡単に計算できるので、まさに教師あり学習に適している。

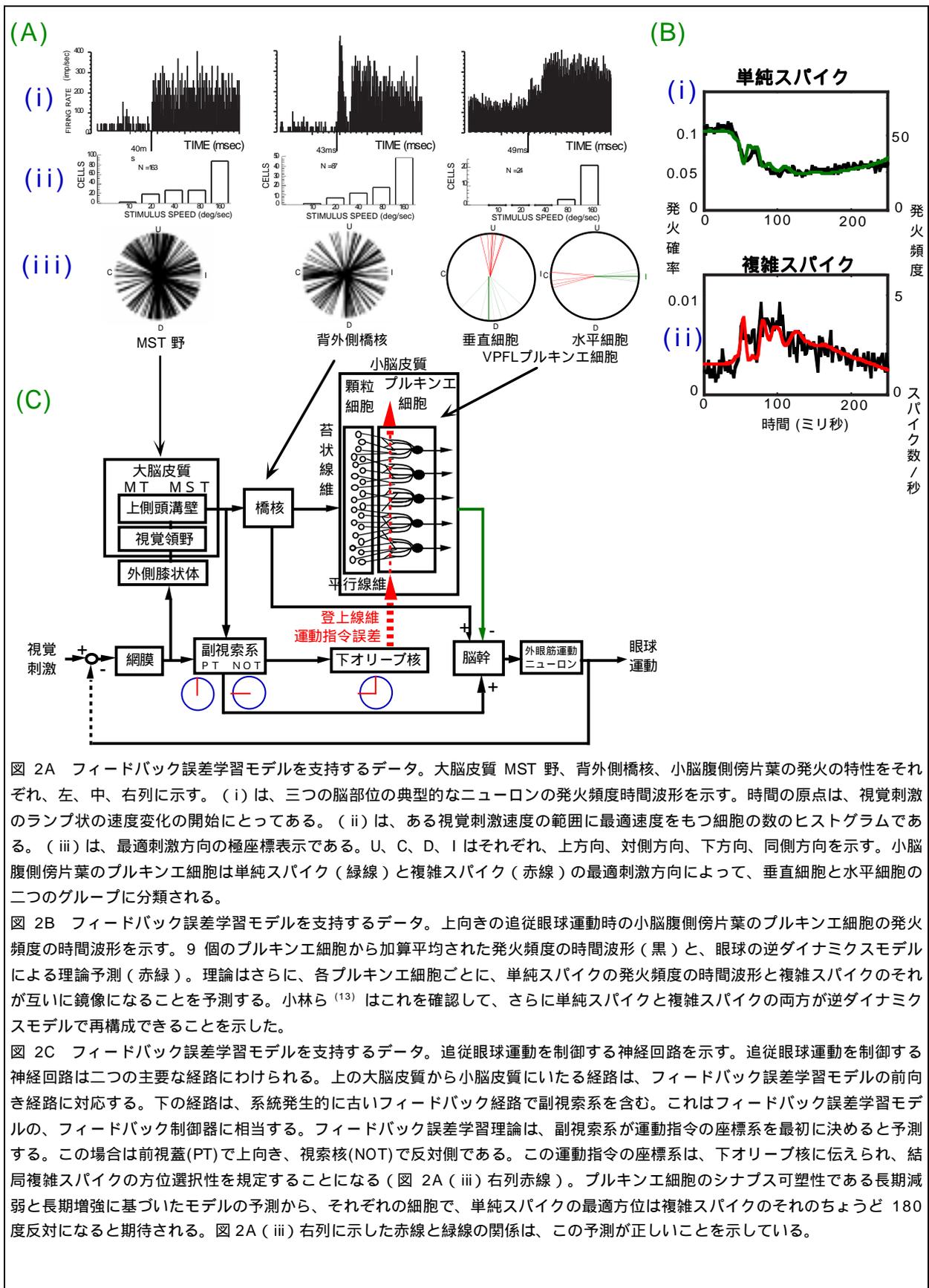


図 2A フィードバック誤差学習モデルを支持するデータ。大脳皮質 MST 野、背外側橋核、小脳腹側傍片葉の発火の特性をそれぞれ、左、中、右列に示す。(i)は、三つの脳部位の典型的なニューロンの発火頻度時間波形を示す。時間の原点は、視覚刺激のランプ状の速度変化の開始にとりである。(ii)は、ある視覚刺激速度の範囲に最適速度をもつ細胞の数のヒストグラムである。(iii)は、最適刺激方向の極座標表示である。U、C、D、Iはそれぞれ、上方向、対側方向、下方向、同側方向を示す。小脳腹側傍片葉のプルキンエ細胞は単純スパイク(緑線)と複雑スパイク(赤線)の最適刺激方向によって、垂直細胞と水平細胞の二つのグループに分類される。

図 2B フィードバック誤差学習モデルを支持するデータ。上向き追従眼球運動時の小脳腹側傍片葉のプルキンエ細胞の発火頻度の時間波形を示す。9個のプルキンエ細胞から加算平均された発火頻度の時間波形(黒)と、眼球の逆ダイナミクスモデルによる理論予測(赤線)。理論はさらに、各プルキンエ細胞ごとに、単純スパイクの発火頻度の時間波形と複雑スパイクのそれが互いに鏡像になることを予測する。小林ら⁽¹³⁾はこれを確認して、さらに単純スパイクと複雑スパイクの両方が逆ダイナミクスモデルで再構成できることを示した。

図 2C フィードバック誤差学習モデルを支持するデータ。追従眼球運動を制御する神経回路を示す。追従眼球運動を制御する神経回路は二つの主要な経路にわけられる。上の大脳皮質から小脳皮質にいたる経路は、フィードバック誤差学習モデルの前向き経路に対応する。下の経路は、系統発生的に古いフィードバック経路で副視索系を含む。これはフィードバック誤差学習モデルの、フィードバック制御器に相当する。フィードバック誤差学習理論は、副視索系が運動指令の座標系を最初に決めると予測する。この場合は前視蓋(PT)で上向き、視索核(NOT)で反対側である。この運動指令の座標系は、下オリーブ核に伝えられ、結局複雑スパイクの方位選択性を規定することになる(図 2A (iii) 右列赤線)。プルキンエ細胞のシナプス可塑性である長期減弱と長期増強に基づいたモデルの予測から、それぞれの細胞で、単純スパイクの最適方位は複雑スパイクのそののちょうど 180 度反対になると期待される。図 2A (iii) 右列に示した赤線と緑線の関係は、この予測が正しいことを示している。

実際に、我々が発展させた理論は、登上線維入力外界のモデリングの対象からのフィードバック情報に基づいた誤差信号を運んでいけば、小脳皮質に、プルキンエ細胞の可塑性に基づいて内部モデルが獲得できることを示した^{(7)~(10)}。この理論についてみていこう。

フィードバック誤差学習

先に述べたとおり、逆モデルは、目標軌道を受け取って、正しい運動指令を出力しなければならない。もし、脳に正しい運動指令を計算する他の部位があれば、それを教師の情報として使って逆モデルを獲得することができる。ところが、そもそも脳に正しい運動指令を計算する機構があったとすれば、それを使って運動制御をおこなえばよいわけで、正しい運動指令を教えてくれる教師が存在するという都合のよい仮定はできない。

つまり、逆モデルを獲得するために脳は、運動をおこなった結果得られる運動軌道の空間での誤差を、運動指令の空間の誤差に変換するという重大な計算論的困難を含む問題を解いていると考えざるをえない。フィードバック誤差学習理論は、登上線維入力、フィードバック運動指令を伝えることによって、そのような運動指令の誤差信号を提供していると提案した⁽⁷⁾⁽⁸⁾⁽¹⁰⁾。この仮説は、神経生理学的研究と詳細な神経回路モデルを組み合わせて、かなりの程度まで証明された^{(11)~(14)}。

サル⁽¹²⁾の追従眼球運動中（大きな視覚パターンを移動させるときの比較的単純な眼球の運動を調べる実験）に、小脳腹側傍片葉のプルキンエ細胞が特徴的なスパイクの列を発生させる。そのスパイク発火頻度変化の時間波形がフィードバック誤差学習理論の予測とよく一致した。これをはじめとして、理論の予測を支持するデータが実にさまざまに、かつ異なる側面に関して得られた（図 2A、図 2B、図 2C）。図 2A、図 2B、図 2C はフィードバック誤差学習モデルを支持するデータで、サル小脳腹側傍片葉と、追従眼球運動に関する、生理学と解剖学データを要約する⁽¹⁴⁾。図 2A と図 2B に示されている神経生理データは小林ら⁽¹³⁾、河野ら⁽¹⁵⁾、竹村ら⁽¹⁶⁾による。図 2A、図 2B、図 2C に示したデータは全体としてフィードバック誤差学習の最も本質的な仮定、登上線維が誤差信号を運動指令の座標系で表現していて、これに誘導されて、単純スパイクの波形が、制御対象の逆ダイナミクスモデルを構成するように、学習で獲得されることを強く支持している。大脳皮質と橋核では、情報がポピュレーション符号化されているのにたいして、小脳皮質の出力、単純スパイクは、情報をきめられた運動指令の座標系で発火頻度の時間波形によって、発火率符号化している。山本ら^{(17)、(18)}はフィードバック誤差学習にもとづくシミュレーションによって、神経符号化の劇的な変化も含めて、図 2A、図 2B、図 2C にまとめられているほぼすべてのデータが、再現できることを示した。

単純な眼球運動について系統発生的に比較的古い小脳部位で検証されたこの理論が、腕の運動にかかわる系統発生的に新しい小脳部位でも正しいことを示唆するデータも得られはじめている⁽¹⁹⁾。

小脳の高次認知機能

小脳の系統発生的に新しい部位は、運動制御ではなく、運動の想像、計画、あるいは運動と直接の関係のない高次認知機能、例えば、パターン認識、心的操作、視覚的注意、視覚運動認知、自閉症、言語、暗算、思考などとかかわることがわかってきた^{(20)~(31)}。代的な研究を表 1 にまとめた。その中でも、と

くに言語とヒト知性にかかわる三つの研究だけを簡単に紹介しよう。

Raichle ら⁽²¹⁾は、名詞から動詞を連想するときに、ともに言語中枢だといわれてきた大脳皮質のブローカ野とウェルニッケ野と同等かそれ以上に右小脳半球が活動することを PET (陽電子放射断層撮影) 計測でみいだした。被験者に与えられた課題は、1.5 秒に 1 個の割合で、誰でも知っているよ

うな英語の名詞を視覚的に呈示されたら、それに対応する適切な動詞をしゃべるといったものであった。例えば“ハンマー”という語を視覚的に呈示されたら、“たたく”と答えるのである。差分画像(目的の課題とコントロールとの脳活動の差を計算する)を作るためのコントロール状態の課題は、みた名詞の復唱であった。

この結果は、コントロール課題で活動する小脳の両側傍虫部は発話運動制御にかかわっているが、右小脳半球は単語の選択という言語課題で活動することを示した。

小脳半球が言語学習にかかわっていることも、関連しておこなわれたつぎのような実験によって示された。上で述べた最初の計測が終わった後、40

個の名詞から成るリストを 15 分間学習し再度同様の計測をおこなうと小脳半球の活動がみられなくなる。しかし、新しいリストに切り替え計測をおこなうと、また強い活動がみられるようになる。この結果をどう解釈したらよいのだろう。

PET や fMRI (機能的磁気共鳴映像法) は、ニューロン活動に伴う血流の増加をみている。登上線維入力による複雑スパイクは平行線維入力による単純スパイクの 100 から 1000 倍の代謝を必要とするので、学習中の PET 計測は誤差信号である登上線維入力を主に観察していると考えられる。我々のフィードバック誤差学習理論にもとづくこの解釈を支持する結果は、後で紹介する今水ら⁽²²⁾の研究からも示される。

マカクサルを用いた生理学的な実験で、小脳が作業記憶や言語機能と関係していることが示されている。例えば、小脳にある歯状核の腹外側部が、空間的作業記憶に役立っている大脳皮質 46 野に視床を介して出力を送っている⁽²⁹⁾。いっぽう、歯状核の背内側部は第一次運動野へ、歯状核の中間部はヒトのブローカ野と相同であると考えられている腹側運動前野に出力を送っている⁽³⁰⁾。

被験者 62 人の小脳の体積と種々の能力の相関を調べた研究もある⁽³¹⁾。人差し指でできるだけ多数回鍵をたたく能力(相関係数 $r = 0.2$ 、有意確率 $p < 0.05$)、言語的記憶能力($r = 0.3$ 、 $p < 0.02$)、一般的知能指数($r = 0.2$ 、 $p < 0.07$)と正の相関があった。それに対して大脳皮質左側頭葉の体積は調べられたど

小脳の高次認知機能への関与

I. 脳活動計測

- | | |
|----------------------|----------------------------|
| (1) 運動の想像 | SPECT Ryding et al. (1993) |
| (2) 名詞からの動詞の連想 | PET Raichle et al. (1994) |
| (3) ペグボードパズル fMRI | Kim et al. (1994) |
| (4) 複数の形の視覚識別 | PET Parsons et al. (1995) |
| (5) 心的回転 | PET Parsons et al. (1995) |
| (6) 皮膚感覚による物体認識 fMRI | Gao et al. (Bower) (1996) |
| (7) 視覚的注意 fMRI | Allen et al. (1997) |

II. 臨床

- | | |
|-------------------|-----------------------------------|
| (1) ハノイの塔 小脳皮質変性症 | Grafman et al. (Hallet M.) (1992) |
| (2) 視覚運動認知小脳皮質変性症 | Nawrot & Rizzo (1995) |
| (3) 自閉症患者 小脳サイズ | Courchesne et al. (1995) |

III. 解剖

- | | |
|--------------|---------------------------|
| (1) 46野 HSVI | Middleton & Strick (1994) |
| (2) IQ 小脳サイズ | Paradiso et al. (1997) |

表 1 ヒト小脳が高次認知機能に重要な役割を果たすことを示す、非侵襲脳活動計測、損傷脳、と解剖学的研究のまとめ。

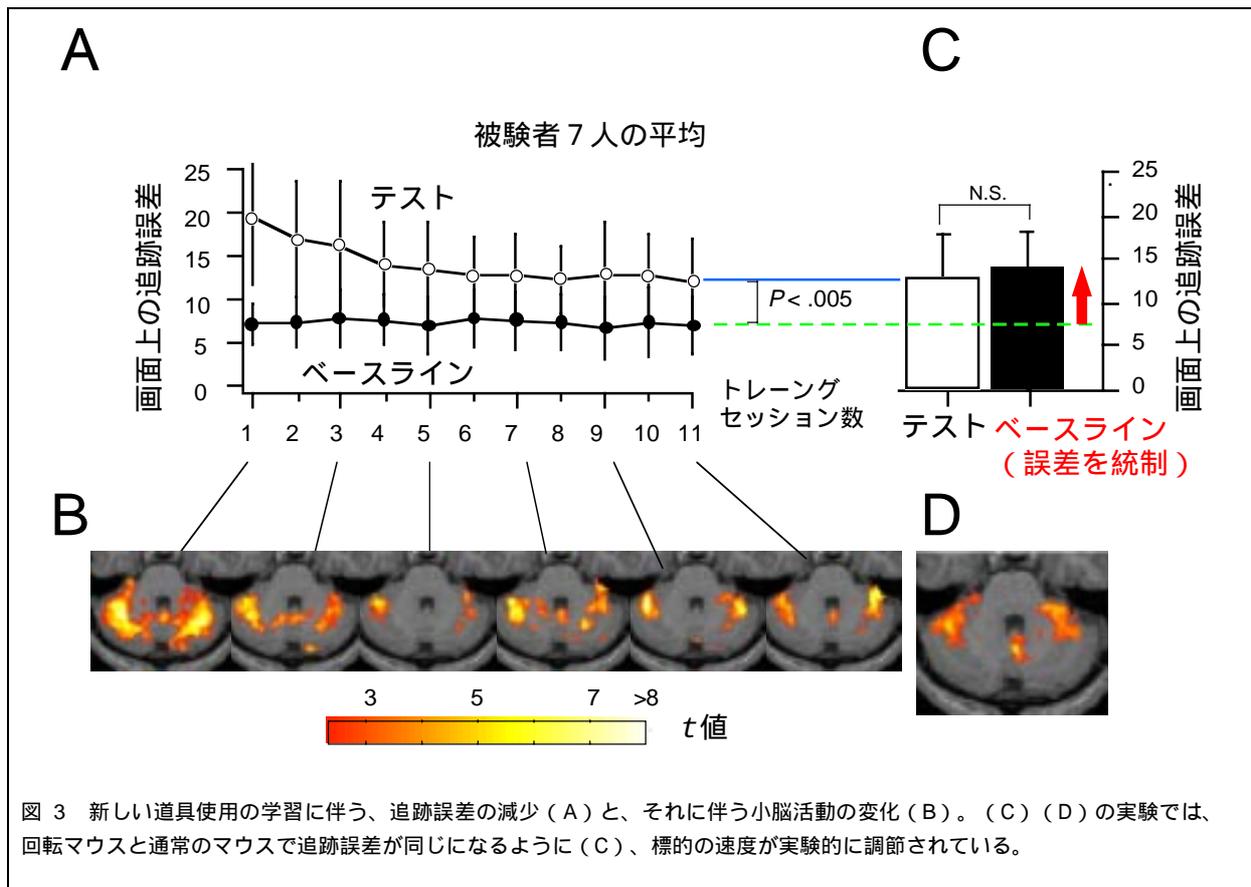
のテストとも有意な相関はなかった。小脳の体積はMRI(磁気共鳴映像法)の画像から自動的に決められ、脳全体の大きさの影響を打ち消すために、脳全体から小脳を差し引いた体積も説明変数の一つに加えられた。

道具の内部モデルの学習

上でも述べたとおり、小脳は実に広範囲な認知活動に役立っているが、一見無関係とも思える多くの機能を説明する統一的な計算理論はあるのだろうか。

今水ら⁽²²⁾は、fMRI法を用いて、ヒト小脳の系統発生的に比較的新しい場所に、道具の内部モデルが、誤差に誘導される教師あり学習で獲得されることを明らかにした。これによって、小脳内部モデル仮説が、運動制御から、ヒト特有の高次認知機能に拡張できることが示唆された。さらに小脳皮質の神経回路構造と、シナプス可塑性は一様であることも考えあわせれば、小脳の系統発生的に新しい部位も、外界のさまざまな対象の内部モデルを学習で獲得していると考えられる。そしてこの場合の外界の対象は、運動制御対象ではなく、例えば道具や、脳の他の部位、他者の脳である。この三つの例では、小脳の機能は、道具使用、思考、コミュニケーションとなる。

今水ら⁽²²⁾は、ヒト被験者が、新しい道具の使用を学習している最中と、学習が終わった後の小脳の活動を fMRI 計測で調べた。ふつうのコンピュータマウスに細工をして、うまく使えるようになるのに 2、



3 時間かかるような回転マウスと積分マウスを作った。回転マウスでは、マウスの動きと、コンピュー

タ画面上のカーソルの動きの間に 120 度の回転変換が入っている。積分マウスでは、マウスの位置が、カーソルの速度を規定する。

課題は、画面上でランダムに 2 次元的な動き（数個の正弦波の重ね合わせで作りだした動き）をす
る標的を、上記の特殊なマウスを操作し、カーソルを動かして、追跡することである。したがって課題
の誤差つまり追跡誤差は、標的とカーソルの距離の時間平均とする。は、11 回のトレーニングセッショ
ンでの追跡誤差と図 3A、奇数回目のセッ
ションで計測された小脳活動図 3B を並べ
て示している。7 人の被験者から得られた
データが平均されている。

図 3A で白丸で示したテストは、回転マ
ウスを使った課題である。いっぽう黒丸
で示したベースラインは、細工をしない
通常マウスを使った課題である。通常マ
ウスについては誤差がセッション数とと
もに変化しないが、回転マウスでは誤差
が訓練とともに減少し、11 セッションめ
ではほぼ定常状態に達していることがわ
かる。

図 3B の fMRI 信号は、テストでベース
ラインより有意に活動が高い場所を統計
的検定の有意さ (t 値) で示してある。学
習のはじめには、広い範囲で活動がみら
れたが、学習の終わりには小さな場所に
だけ活動が残っている。

運動課題や認知課題の学習に伴って小
脳の活動が顕著に減少することは、過去
の非侵襲脳活動計測でも観測されていた。
これらのデータをもとに、小脳は学習の
初期には重要であるが、学習が完了した
ときには記憶の痕跡は（つまり内部モデ
ルも）小脳にはないという解釈がしばし
ばなされてきた。しかし、図 3C と D に
示すデータは、この考え方を否定した。

図 3C と D に示した実験では、ベース
ライン条件で、標的の速度を増加させ、
テストとベースラインでの追跡誤差を等しくした。この条件では、視覚刺激速度、運動の大きさ、注意、

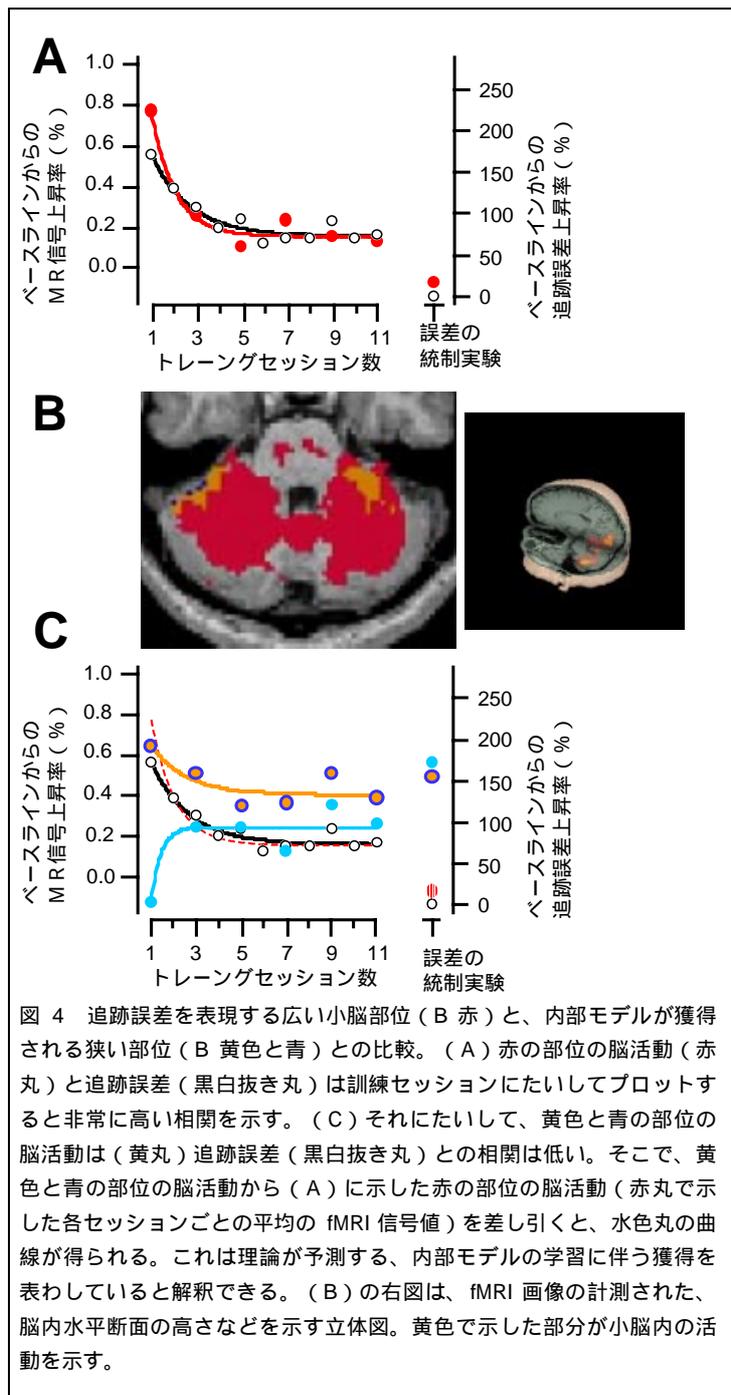


図 4 追跡誤差を表現する広い小脳部位 (B 赤) と、内部モデルが獲得される狭い部位 (B 黄色と青) との比較。(A) 赤の部位の脳活動 (赤丸) と追跡誤差 (黒白抜き丸) は訓練セッションにたいしてプロットすると非常に高い相関を示す。(C) それにたいして、黄色と青の部位の脳活動は (黄丸) 追跡誤差 (黒白抜き丸) との相関は低い。そこで、黄色と青の部位の脳活動から (A) に示した赤の部位の脳活動 (赤丸で示した各セッションごとの平均の fMRI 信号値) を差し引くと、水色丸の曲線が得られる。これは理論が予測する、内部モデルの学習に伴う獲得を表わしていると解釈できる。(B) の右図は、fMRI 画像の計測された、脳内水平断面の高さなどを示す立体図。黄色で示した部分が小脳内の活動を示す。

努力などすべての要因がベースライン条件のほうが大きい。それにもかかわらず、テスト条件の方で小脳活動が大きい部位が図 3) のようにみつかったのである。したがってテスト条件でより大きくなる小脳の活動は、新しい道具の使用法の記憶、さらには、回転マウスの内部モデル以外として解釈することはできない。

学習の最初に広く活動する部位 (図 4B 赤) の信号値を空間的に平均して、セッション数に対してプロットすると、追跡誤差とほとんど同じ時間経過をたどることがわかる (図 4A)。いっぽう最後まで活動が残る部位 (図 4B 黄色と青) では、信号値は追跡誤差のように減少せず、追跡誤差に対応する赤の部位の活動を差し引くと、図 4C の水色で示したように学習とともに増加する曲線が得られる。これは学習とともに徐々に獲得される内部モデルの活動に対応していると考えられる。

この結果から、系統発生的に新しい小脳部位での道具使用という認知的な課題でも、フィードバックとなる。つまり、赤の部位での信号は、登上線維が表現している運動誤差を表わしている。いっぽう、

黄色と青の部位では、学習によって内部モデルが徐々に獲得され、その活動が学習が完了しても観測される。

さらに玉田ら⁽³²⁾は、大脳と小脳の活動の左右差が反対称になることを利用して、最後に活動の残る小脳部位が、大脳皮質運動前野腹側部 (三角部と弁蓋部) と機能的に結合していることを示している (図 5)。この大脳部位の左側は、言語野であるブローカ野を含んでおり、道具使用と言語の神経機構が重なっていることを示している。

さらに今水ら⁽³³⁾は、複数の道具にたいして、小脳の異なる部位が活動することも示している。また玉田ら⁽³⁴⁾は、把持力負荷力結合という行動課題を用いて、物体の順モデルが小脳内に存在することを示している。

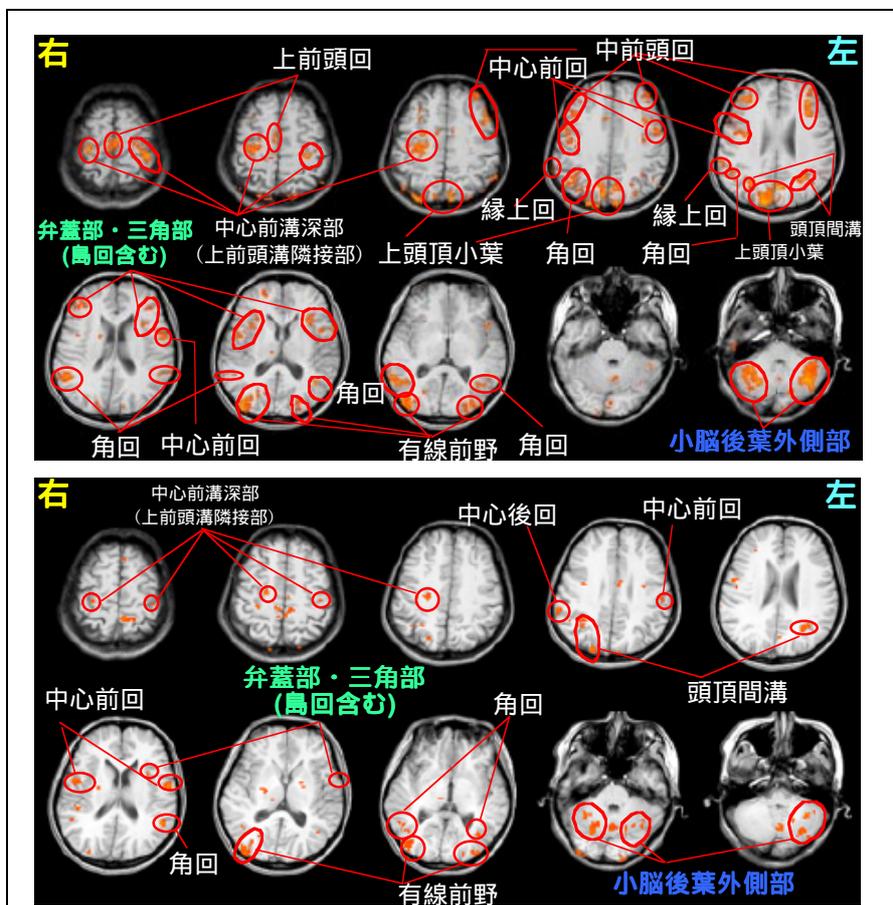


図 5 道具使用の学習に伴う、大脳小脳機能連関図。代表的なひとりの被験者における学習初期 (上) と学習後期 (下) の脳活動。色をつけた部分が、ベースライン条件よりテスト条件で有意に高い ($P < 0.0001$) 信号値を示した領域。図の左上が頭頂、左下が小脳下部の水平断面。大脳皮質のさまざまな脳部位が、道具使用の学習に伴って活動することがわかる。これらの部位のうち、大脳と小脳では対側に結合があるという事実に基づいて、小脳と機能的に結合されている場所を探ると、ブローカ野だけが残った。

これらの実験データは、多重順逆モデルを強く支持するデータとなっている。

分裂病とくすぐりの予測

他人に身体を触られるとくすぐりたいのに、自分で触ってもくすぐったくないのはどうしてだろう。小脳の中の順モデルが、自分自身の運動のときは、運動指令の遠心性コピーから、触覚を予測し、それが実際の感覚フィードバックから差し引かれるので、くすぐったくないのである⁽³⁵⁾。

分裂病の患者では、自分で自分をくすぐっても、十分くすぐりたいと感じる。このとき、小脳の順モデルの活動が、正常者と異なることもわかっている。分裂病の患者が示す症状の多く、特に幻聴などの運動制御の妄想が、自分自身の運動指令からひきおこされる感覚信号の変化と、外界の状況の変化によってひきおこされる感覚信号の変化とを区別できないことによると解釈できる。つまり、分裂病が、小脳内の順モデルが機能しないことによる病態として理解できるのである。

運動制御から言語にいたるまで、小脳が同じ計算原理を用いていることを説明した。小脳皮質の神経回路構造は一樣であるから、計算原理も系統発生的に新しい部分と古い部分で共通である。系統発生的に古い小脳では、運動制御対象（身体の一部）の逆モデルが、シナプス可塑性に基づく運動学習で獲得されることが示された。系統発生的に新しい小脳部位でも、言語の連想や新しい道具の使用など、認知的な課題で、誤差に誘導される学習が生じていることが、PET や fMRI といった非侵襲脳活動計測からわかった。

ヒト小脳には複数の道具に対応して、複数の内部モデルが存在する。ヒト小脳の順モデルの異常は精神分裂病の原因となる。小脳皮質のシナプス可塑性と、神経回路をこれらのデータに基づいて計算論的に解釈すれば、小脳皮質は、運動制御から高次認知機能までの異なる機能にかかわらず、入力を出力に変換する非線形写像（内部モデル）を、教師あり学習で獲得するといえる。

文献

- (1) Middleton FA, Strick PL: The cerebellum: an overview. *Trends Cog. Sci.*, 2, 305-306 (1998) and also in *Trends Neurosci.*, 21, 367-369 (1998)
- (2) Ito M: *The Cerebellum and Neural Control*, Raven Press (1984)
- (3) Marr D: A theory of cerebellar cortex. *J. Physiol.*, 202, 437-470 (1969)
- (4) Albus JS: A theory of cerebellar functions. *Math Biosci.*, 10, 25-61 (1971)
- (5) Ito M: Neurophysiological aspects of the cerebellar motor control system. *Int. J. Neurol.*, 7, 162-176 (1970)
- (6) Gomi H, Kawato M: Equilibrium-point control hypothesis examined by measured arm-stiffness during multi-joint movement. *Science*, 272, 117-120 (1996)
- (7) Kawato M, Furukawa K, Suzuki R: A hierarchical neural-network model for control and learning of voluntary movement. *Biol Cybern.*, 57, 169-185 (1987)
- (8) Kawato M, Gomi H: Computational models of cerebellar motor learning. *Trends in Neurosci.*, 16, 177-178 (1993)

- (9) Miall RC, Weir DJ, Wolpert D, Stein JF: Is the cerebellum a Smith Predictor? *J Motor Behavior*, 25, 203-216 (1993)
- (10) Wolpert D, Miall C, Kawato M: Internal models in the cerebellum. *Trends in Cog. Sci.*, 2, 338-347 (1998)
- (11) Shidara M, Kawano K, Gomi H, Kawato M: Inverse-dynamics model eye movement control by Purkinje cells in the cerebellum. *Nature*, 365, 50-52 (1993)
- (12) Gomi H, Shidara M, Takemura A, Inoue Y, Kawano K, Kawato M: Temporal firing patterns of Purkinje cells in the cerebellar ventral paraflocculus during ocular following responses in monkeys. I. simple spikes. *J. Neurophysiol.*, 80, 818-831 (1998)
- (13) Kobayashi Y, Kawano K, Takemura A, Inoue Y, Kitama T, Gomi H, Kawato M: Temporal firing patterns of Purkinje cells in the cerebellar ventral paraflocculus during ocular following responses in monkeys. II. complex spikes. *J. Neurophysiol.*, 80, 832-848 (1998)
- (14) Kawato M: Internal models for motor control and trajectory planning. *Curr. Opin. Neurobiol.* 9, 718-727 (1999)
- (15) Kawano K, Takemura A, Inoue Y, Kitama T, Kobayashi Y, Mustari MJ: Visual inputs to cerebellar ventral paraflocculus during ocular following responses. *Prog. Brain Res.*, 112, 415-422 (1996)
- (16) 竹村文、井上由香、五味裕章、川人光男、河野憲二: 追従眼球運動時の3つの脳内領域のニューロン活動の解析. 電子情報通信学会技術研究報告, NC99-22, 77-84 (1999)
- (17) Yamamoto K, Kobayashi Y, Takemura A, Kawano K, Kawato M: A mathematical model that reproduces vertical ocular following responses from visual stimuli. *Neurosci. Res.*, 29, 161-169 (1997)
- (18) 山本憲司、小林康、竹村文、河野憲二、川人光男: 垂直方向追従眼球運動の適応シミュレーション: 小脳皮質の計算モデルは運動学習を再現できるか? 電子情報通信学会技術研究報告, NC97-131, 229-236 (1998)
- (19) Kitazawa S, Kimura T, Yin P: Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature*, 392, 494-497 (1998)
- (20) Ryding E, Decety J, Sjöholm H, Stenberg G, Ingvar DH: Motor imagery activates the cerebellum regionally. A SPECT rCBF study with 99mTc-HMPAO. *Brain Res. Cogn. Brain Res.*, 1, 94-99 (1993)
- (21) Raichle ME, Fiez JA, Videen TO, MacLeod AM, Pardo JV, Fox PT, Petersen SE: Practice-related changes in human brain functional anatomy during nonmotor learning. *Cereb. Cortex*, 4, 8-26 (1994)
- (22) Imamizu H, Miyauchi S, Tamada T, Sasaki Y, Takino R, Pütz B, Yoshioka T, Kawato M: Human cerebellar activity reflecting an acquired internal model of a novel tool. *Nature*, 403, 192-195 (2000)

- (23) Kim SG, Ugurbil K, Strick PL: Activation of a cerebellar output nucleus during cognitive processing. *Science*, 265, 949-951 (1994)
- (24) Parsons LM, Fox PT, Downs JH, Glass T, Hirsch TB, Martin CC, Jerabek PA, Lancaster JL: Use of implicit motor imagery for visual shape discrimination as revealed by PET. *Nature*, 375, 54-58 (1995)
- (25) Gao JH, Parsons LM, Bower JM, Xiong J, Li J, Fox PT: Cerebellum implicated in sensory acquisition and discrimination rather than motor control. *Science*, 272, 545-547 (1996)
- (26) Allen G, Buxton RB, Wong EC, Courchesne E: Attentional activation of the cerebellum independent of motor involvement. *Science*, 275, 1940-1943 (1997)
- (27) Grafman J, Litvan I, Massaquoi S, Stewart M, Sirigu A, Hallett M: Cognitive planning deficit in patients with cerebellar atrophy. *Neurology*, 42, 1493-1496 (1992)
- (28) Courchesne E, Akshoomoff NA, Townsend J, Saitoh O: A model system for the study of attention and the cerebellum: infantile autism. *Electroencephalogr. Clin. Neurophysiol. Suppl.*, 44, 315-325 (1995)
- (29) Middleton FA, Strick PL: Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science*, 266, 458-461 (1994)
- (30) Strick PL, Hoover JE, Mushiakhe H: Evidence for "output channels" in the basal ganglia and cerebellum. in "Role of the Cerebellum and Basal Ganglia in Voluntary Movement", Mano N, Hamada I, DeLong MR (Eds), Excerpta Medica, Tokyo, pp.171-180 (1993)
- (31) Paradiso S, Andreasen NC, O'Leary DS, Arndt S, Robinson RG: Cerebellar size and cognition: correlations with IQ, verbal memory and motor dexterity. *Neuropsychiatry Neuropsychol Behav. Neurol.* , 10, 1-8 (1997)
- (32) Tamada T, Miyauchi S, Imamizu H, Yoshioka T, Kawato M: Cerebro-cerebellar functional connectivity revealed by the laterality index in tool-use learning. *NeuroReport*, 10, 325-331 (1999)
- (33) Imamizu H, Miyauchi S, Tamada T, Sasaki Y, Takino R, P tz B, Yoshioka T, Kawato M: Multiple representations for visuomotor learning in the cerebellum: A functional MRI study. *NeuroImage*, 7, S819 (1998)
- (34) Tamada T, Miyauchi S, Imamizu H, Yoshioka T, Kawato M: Cerebellar activation in grip force adjustments during transporting an object held in a precision grip. *Abst. Ann. Soc. Neurosci.*, 25, 1894 (1999)
- (35) Blakemore SJ, Wolpert D, Frith CD: Central cancellation of self-produced tickle sensation. *Nature Neurosci.*, 1, 635-640 (1998)

人間の脳に獲得される内部モデル：脳機能イメージングによる検証

今水 寛

科学技術振興事業団川人学習動態脳プロジェクト

1. はじめに

従来、小脳は運動制御のための神経機構であると言われてきた。しかし、人間の脳活動を非侵襲的に計測できるようになり、そう単純には言い切れなくなった。言葉を連想したり、パズルの解法を考えるなど、被験者が身体運動を伴わない認知的な課題を行っているときでも、小脳の活動レベルが上昇することが解ってきたからである。どのような課題を行わせると小脳活動が上昇するかという経験的知見は、次々に蓄積されてきた（総説として、Thach 1996; Desmond and Fiez 1998 など）。しかし一方で、活動の意味については、多くの謎が残っている。人間以外の動物で詳細に調べられてきた神経生理学的知見と、それに基づく理論的なモデルとどのように結びつくのか？ 小脳はどのようなメカニズムで認知機能に役立っているのか？

筆者らはこの謎を解明するために、計算論的神経科学の枠組みに基づき、小脳で計測される活動が何を反映しているかについての仮説を立てた。さらに、人間の被験者が新しい道具の使い方を学習しているときの小脳活動を計測し、仮説の妥当性を検証した。

2. これまでの研究

2-1. 内部モデル：理論的枠組み

中枢神経系における運動制御の仕組みを理解するうえで、内部モデルという概念が重要になりつつある（総説として Kawato 1999）。随意運動における内部モデルとは、「ある運動司令がどのような動作を引き起こすか」（順モデル）、あるいは「ある動作を行いたいときに、どのような運動司令を出せばよいか」（逆モデル）という対応関係の内部表現であると考えられる。このような対応関係が、運動開始前から解っていれば、感覚フィードバックに頼らなくても、速く正確な運動制御が行える。

内部モデルが、どのような仕組みで学習されるかということに関して、川人ら（Kawato, Furukawa and Suzuki 1987; Kawato and Gomi 1992）は「フィードバック誤差学習スキーム」というモデルを提唱した（図 1 A）。運動のはじめは、感覚フィードバックに頼ってぎこちない運動をしている（aのループ）が、フィードバック制御システムの出力を「誤差信号」として、内部モデルをトレーニングする（bの矢印）。繰り返し練習することで、内部モデルは、意図した動作から、それを実現するための運動司令への変換ができる

ようになり (c のルート), 感覚フィードバックに依存しなくても, 速くて正確な制御ができるようになる。

2-2. 内部モデル：神経生理学の実体

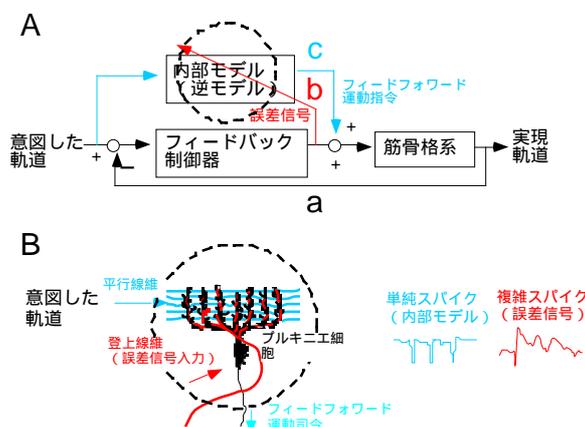


図1 フィードバック誤差学習のスキーマ

上記のスキーマでは, 小脳皮質のプルキニエ細胞が内部モデルとして重要な役割を果たしている (図1B). プルキニエ細胞には, 主に平行線維入力と登上線維入力という2種類の入力がある. 例えば, 随意運動の逆モデルを学習する場合, このスキーマでは, 平行線維入力は「意図した軌道 (動作)」を表現し, 登上線維入力は「意図した軌道」と「実現した軌道」のずれ (誤差信号) を表現していると考えられる. プル

キニエ細胞からの出力は「運動司令」である. 学習のはじめ, 誤差が大きいつきには, 登上線維を伝って誤差信号が盛んに入力される. 誤差信号はプルキニエ細胞のシナプスの伝達効率を変化させ, 最終的には小脳皮質において, 「意図した軌道」からそれを実現させる「適切な運動司令」への変換ができるようになる.

登上線維入力によって複雑スパイクが, 平行線維入力によって単純スパイクが引き起こされることが知られている (図1B左). 上記のスキーマは, 複雑スパイクは誤差信号を反映し, 単純スパイクは内部モデルの出力である運動司令を反映すると予測する. 実際, サルを用いた神経生理学的な実験で, この予測が確かめられている (複雑スパイクに関しては, Kobayashi, et al. 1998; Kitazawa, Kimura and Yin 1998; 単純スパイクに関しては, Shidara, et al. 1993; Gomi, et al. 1998).

2-3. 内部モデル理論と非侵襲計測脳活動

以上のような内部モデル理論は, 学習中の小脳において2つのタイプの活動が計測できると予測する. 第一の活動は誤差信号を反映し (図2A赤い曲線), 第二の活動は学習によって獲得される内部モデルの活動を反映する (図2B水色の曲線). 誤差を反映する活動は, 学習の初期に誤差が大きいつき, 広い範囲で信号値 (PETやfMRIの測定値) の上昇として確認できるはずである. 一方, 内部モデルの活動は, 学習が進むにつれて上昇すると考えられる.

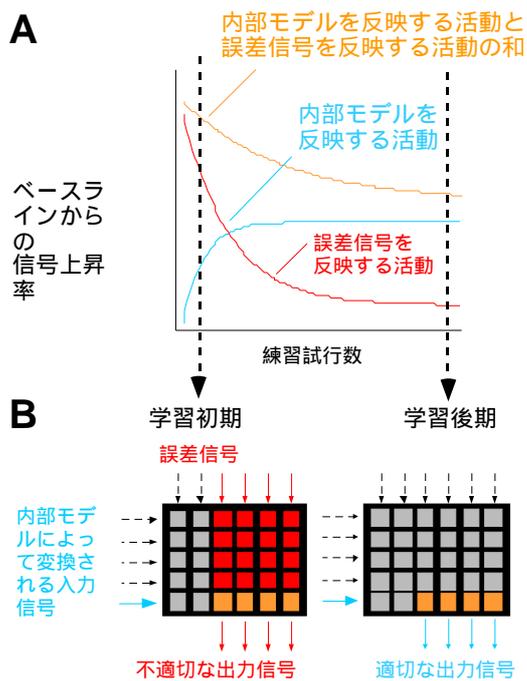


図2 小脳活動の理論的予測

これまでの脳活動非侵襲計測実験では、ある新しい課題を学習するとき、初期には小脳の広い範囲で強い活動の上昇が見られ、学習が進むに従って、活動の範囲も強度も減少すると言われてきた。Raichleら（1994）は名詞から動詞を連想する課題で、Flamentら（1996）はジョイスティックを用いた視覚運動課題で、小脳活動が学習とともに減少する様子を報告している。内部モデル理論の枠組みで考えると、これらの先行研究は、主に第一の活動（誤差信号を反映する活動）を見ていると思われる。しかし、これらの結果は、「小脳皮質は学習の初期にのみ重要な役割を果たし、練習によって獲得された記憶は、脳の別な場所に蓄えられる」という説（例えば Raymond, et al. 1996）を支持し、小脳が記憶の座である

（例えば Ito and Kano 1982）という説を否定しているように考えられてきた。本当であろうか？

内部モデル理論の妥当性を検証するには、第二のタイプの活動（内部モデルを反映する活動）が計測できるかどうかを確認する必要がある。しかし、単純なイメージングの方法では計測できない。なぜならば、誤差の情報も内部モデルが獲得される場所（図2 Bのオレンジの領域）にも入力されるはずであり、観測できるのは第一の活動と第二の活動の和（図2 Aのオレンジの曲線）である。そこで、筆者らは、テスト条件とベースライン条件で、第一の活動レベルが同じになるような実験的操作を行い、第二の活動を計測することを試みた。

3. 最近の研究

3-1. 学習に伴う小脳活動の変化

具体的な実験としては、被験者が新しい道具の使い方を学習しているときの小脳活動を計測した（Imamizu et al., 2000）。基本的な実験課題はコンピュータマウスを操作して、画面上をランダムに動き回るターゲットを追跡することであった（トラッキング課題）。ただし、テスト条件では、画面上のカーソルが120°回転した位置に表示された（回転

マウス)。ベースライン条件では、そのような変換は行わなかった（通常マウス。被験者はまず11セッションのトレーニングを受けた。1セッションは9分23秒で、35.2秒ごとに、テスト条件とベースライン条件が交替した。トレーニング中、奇数番目のセッションでは、fMRIで小脳活動を記録した。

行動レベルのデータとしては、ターゲットカーソルの間の距離を4.4秒間（小脳活動のスキャン間隔と同じ時間）積算した値を記録した。この値は「トラッキング誤差」であり、完璧にターゲットを追跡できれば、0になる。図3Aは、トレーニング中にトラッキング誤差がどのように変化したかを示している。ベースライン条件では、ほぼ一定であったのに対し、テスト条件では次第に減少し、学習が進んでいることを示唆している。図3Bは、同時に計測した小脳活動を統計解析した結果である。テスト条件で有意に信号値が上昇した場所を赤—黄色で示している。学習の始めは、小脳の広い範囲で信号値の増加が見られたが、学習が進むにつれて、活動領域は限定された。

3-2. 活動領域は消えるのか？

ここまでは、学習とともに小脳活動が減少するという先行研究の結果と一致している。問題は、テスト条件とベースライン条件で誤差を同じにしても（=誤差を反映する活動をベースラインとして、さし引いても）活動は残るかということである。

筆者らは次のような方法で誤差の統制実験を行った。まず、ベースライン条件でのトラッキング誤差は、ターゲット速度に比例することが経験的に解っていた。そこで、通常のマウスを用いて、様々なターゲットの速度でトラッキングを行ってもらい、被験者ごとに、ターゲット速度とトラッキング誤差の関係性を推定した。これによって、先行するテスト条件のトラッキング誤差に応じて、ベースライン条件でのターゲット速度を調節し、両

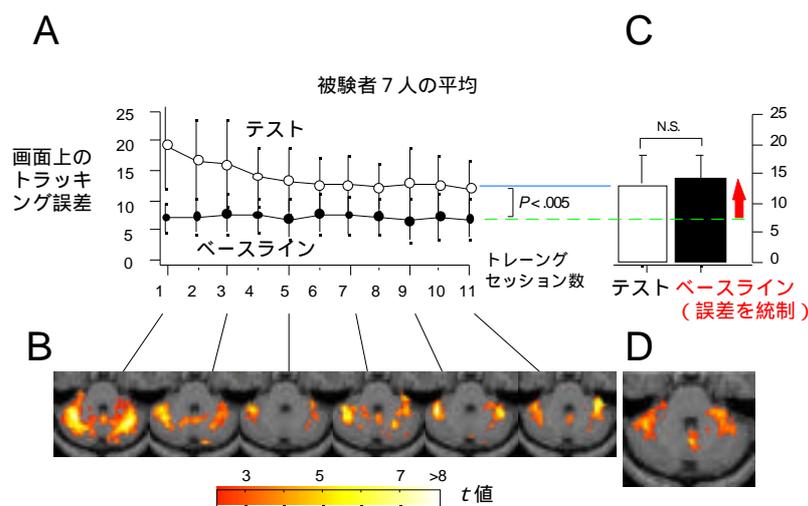


図3 学習に伴う小脳活動の変化

者の誤差を同じにすることができる。

図 3 C は、誤差の統制実験におけるトラッキング誤差を、図 3 D は、同時に計測した脳活動を示す。テスト条件とベースライン条件でトラッキング誤差に有意な差がないにもかかわらず、テスト条件で有意に信号値が上昇する小脳領域が存在していた。この活動は、誤差の要因では説明できない。また、ベースライン条件でのターゲット速度の方が速く（平均 2.71 倍）、マウスの移動量（手の動き）・目の動き・視覚刺激の速さ・注意など、被験者に対する認知的・運動的な負荷は、ベースライン条件の方が大きかった。従って、この活動は、純粹に回転マウスの操作に必要な内部モデルの活動を反映していると考えられる。

3-3. 領域ごとの活動の時間変化

内部モデルを反映する活動領域と、誤差を反映する活動領域を調べた。テスト条件で 1、ベースライン条件で 0 となるステップ関数と、上記の誤差の統制実験で計測した小脳活動

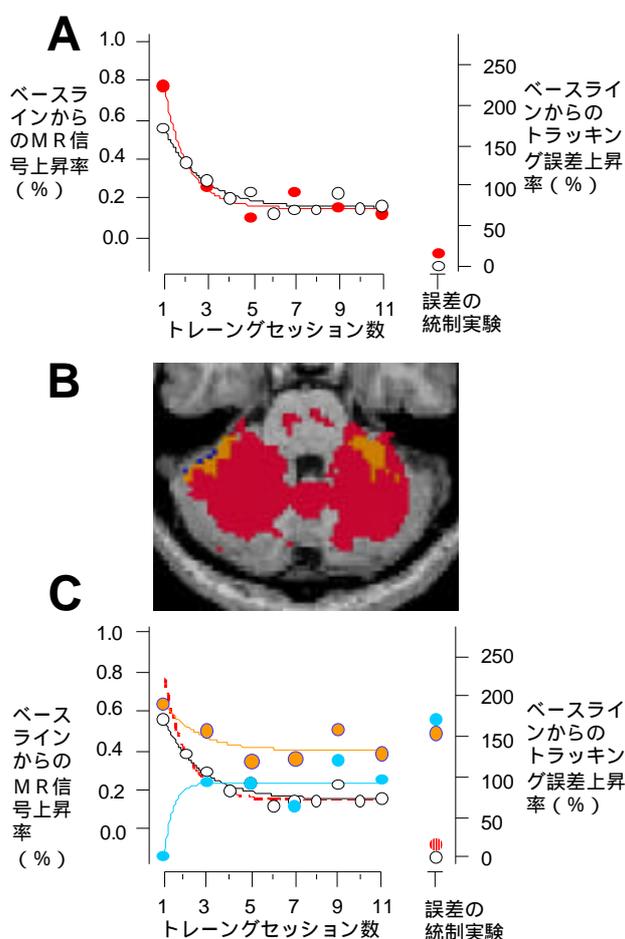


図 4 誤差を反映する活動と内部モデルを反映する活動

の間で回帰分析を行い、回帰係数が有意に 0 よりも大きい領域 ($p < .05$, 多重比較の補正あり) を、「内部モデルを反映する活動領域 (図 4 B の赤い領域)」とした。また、トレーニング中に計測した小脳活動とトラッキング誤差の間で回帰分析を行い、回帰係数が有意に 0 よりも大きい領域を「誤差を反映する活動領域 (図 4 B の青い領域; オレンジの領域は、赤と青の共通領域)」とした。

図 4 A・C は、それぞれの領域における信号値の上昇率 (テスト条件において、ベースライン条件から何%上昇したか) を示している。赤・オレンジの領域では、学習とともに信号値が急速に低下していた (図 4 A)。信号値の上昇率と、トラッキング誤差の上昇率の間で相関を調べると、有意な相関が見られた ($r^2 = .082, p < .005$)。一

方、青・オレンジの領域では、赤い領域ほど低下していなかった（図4C）. トラッキング誤差との相関も赤い領域ほど高くはなく、有意ではなかった（ $r^2=.025$ ）. トレーニング期間全体で、赤・オレンジ領域と青・オレンジ領域の信号上昇率の間で、有意差が見られた（ $p<.05$ ）. これは、青・オレンジ領域の信号値は、誤差だけでは説明できない成分を含んでいることを示唆している. 赤い曲線からオレンジの曲線をさし引くことで、その成分が求められる（図4Cの水色の曲線）が、学習開始とともに上昇している様子がわかる. この成分は、学習によって獲得された内部モデルの活動を反映していると考えられる.

3-4. 2種類の活動がfMRI信号に反映されるメカニズム

【誤差を反映する活動】

誤差を反映する活動は、非常に強い信号値の上昇として観察できた. 内部モデル理論では、登上線維入力によって引き起こされる複雑スパイクが、誤差の情報を反映すると考えられる. 神経生理学の研究（Hockberger, Tseng and Connor 1989, Hounsgaard and Yamamoto 1979）は、小脳で最もエネルギーを消費するのは、複雑スパイク発火後のイオンの平衡状態の回復であると示唆している. 従って、誤差を反映する活動が強い信号値の上昇をもたらすことは、理論的にも生理学的にも矛盾しない.

【内部モデルを反映する活動】

一方、内部モデルの活動や学習過程が、fMRIの増加として計測できることは、これまでの神経生理学的常識とは矛盾するように思える. プルキニエ細胞のシナプス伝達効率を変化させるメカニズムとしては、LTD（Long-term depression; Ito, Sakurai and Tongroach 1982）が良く知られているが、LTDはシナプス伝達効率を減少させ、fMRI信号を減少させる方向に働くと考えられるからである.

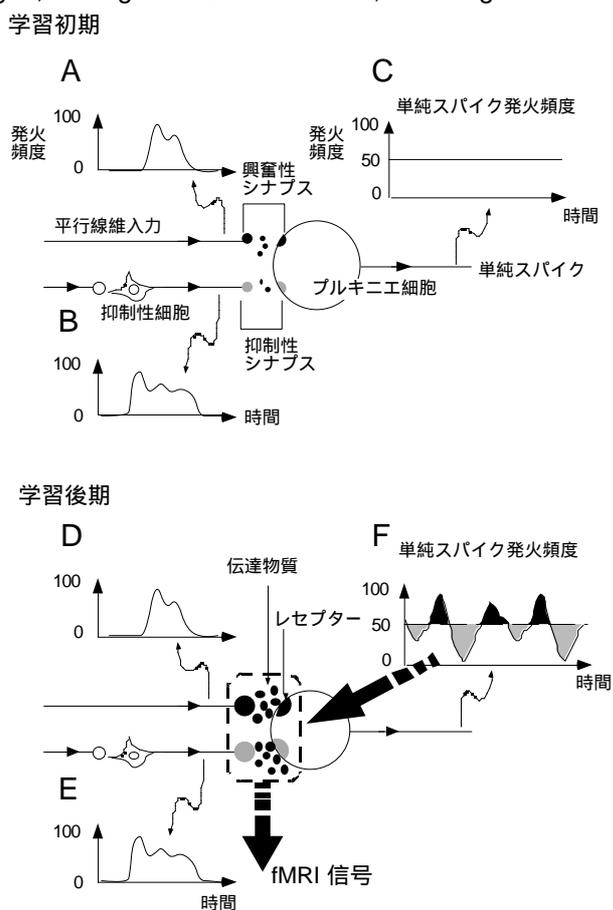


図5 内部モデルを反映する活動がfMRIに反映されるメカニズム

LTD に関連するさまざまな化学反応の中には、信号増加の原因となるものも考えられる (Ito 2000)。例えば、酸化窒素の増加は、毛細血管を広げて局所血流量を増加させる。しかし、今回の実験では、テスト条件とベースライン条件が 35.2 秒という短い時間で交替したので、そのような化学反応が短時間で急激に増加・減少するとは考えにくい。また、LTD に関連する化学反応は、学習初期の広い範囲の活動は説明できても、学習後に誤差のレベルを揃えた実験で、テスト条件で有意に信号が上昇したという結果は説明できない。この場合、テスト条件でもベースライン条件でも、同じ大きさの誤差信号が小脳皮質に入力され、LTD は同様に生じていたと考えられるからである。

内部モデルの活動が、信号を増加させるメカニズムとしては以下のようなことが考えられる。プルキニエ細胞は、平行線維入力の時間波形 (図 5 D, E) を、単純スパイクの発火頻度波形 (F) に適切に変換することを学習する。例えば、プルキニエ細胞が運動制御のための逆モデルとして機能するとき (Kawato and Gomi 1992; 図 1 参照)、平行線維入力は意図した軌道に、単純スパイクの発火頻度は、その軌道を実現するフィードフォワード運動司令に相当する。平行線維 / プルキニエ細胞シナプスには、興奮性・抑制性のシナプスが混在していること、平行線維入力からの情報は、発火頻度の増加で表現されていることから、学習初期には、さまざまな時間波形の平行線維入力は互いに相殺され、単純スパイクの発火頻度波形は平坦であると考えられる (図 5 C)。しかし、単純スパイクの発火頻度波形が、フィードフォワード運動司令として機能するためには、単純スパイクの発火頻度が時間とともに増加・減少する (モジュレーションがおきる) 必要がある。これは、プルキニエ細胞のシナプス伝達効率が変化することで、おきると考えられる。伝達効率を変化させるメカニズムとしては、LTD 以外にも、potentiation (Sakurai 1987) や rebound potentiation (Kano, et al. 1992) が知られている。発火頻度の増加 (図 5 F の黒い部分) には、potentiation が必要である。追従眼球運動の適応を説明するためには、LTD と potentiation の両方が必要であることが、最近の研究 (山本, 小林, 竹村, 河野, 川人, 1998) で明らかになった。potentiation や rebound potentiation によるシナプス伝達効率の増加は、代謝活動の増加を引き起こし、fMRI の信号増加としてとらえられるはずである。

4. おわりに

本稿では第一に、小脳における内部モデル理論と、理論がどのような非侵襲計測脳活動を予測するかについて述べた。第二に、新奇な道具の使い方を学習するときのヒト小脳活動を計測した研究について述べた。この研究では、理論の予測通り、誤差を反映する活動と内部モデルを反映する活動を確認した。第三に、計測された小脳活動の神経生理学的意

味について検討した。

小脳皮質の神経回路構造は一様であるから，計算原理はどの部分でも共通であると考えられる．一方，トレーサーを用いた解剖学的研究技術の進歩は，前頭葉，頭頂葉，後頭葉の一部を含む広範な領域から小脳への投射と，小脳から前頭連合野の広範な領域への投射を明らかにした．従って，小脳は内部モデルを獲得することで，運動制御に限らず，言語や思考を始めとする高次認知脳機能においても，速く正確な情報処理を可能にしていると考えられる．

高次脳機能を支える内部モデルとして機能するためには，複数の内部モデルを混乱なく学習したり，適切な内部モデルを選択する必要が生じる（モジュール性）．また，ある内部モデルが他の内部モデルを要素として参照する必要も生じる（埋め込みによる階層構造）．最近の計算論的研究では，小脳とその周辺の神経回路で，内部モデルのモジュール的な学習と選択（Wolpert and Kawato 1998 など）や埋め込みによる階層構造（川人，1997）を実現できることが明らかにされている．

本稿で紹介した一連の研究は，回転マウスの使い方を学習する時の脳活動という，特殊な事例に基づくことは否定できない．しかし，小脳活動の事例報告をひとつ増やしただけではない．人間以外の動物で詳細に調べられてきた神経生理学的知見と，それに基づく計算理論の立場から，fMRIで計測した小脳活動の意味を詳細に解明した初めての研究である．

文献

- Desmond JE, Fiez JA : Neuroimaging studies of the cerebellum: language, learning and memory. Trends in Cognitive Science 2(9):355-362, 1998
- Flament D, Ellermann JM, Kim SG, Ugurbil K, Ebner TJ : Functional magnetic resonance imaging of cerebellar activation during the learning of a visuomotor dissociation task. Human Brain Mapping, 4(3) :210-226, 1996
- Gomi H, Shidara M, Takemura A, Inoue Y, Kawano K, Kawato M : Temporal firing patterns of purkinje cells in the cerebellar ventral paraflocculus during ocular following responses in monkeys I. Simple spikes. J Neurophysiol 80(2): 818-31, 1998
- Hockberger PE, Tseng HY, Connor JA : Fura-2 measurements of cultured rat Purkinje neurons show dendritic localization of Ca²⁺ influx. J Neurosci 9(7):2272-84, 1989
- Hounsgaard J, Yamamoto C : Dendritic spikes in Purkinje cells of the guinea pig cerebellum studied in vitro. Exp Brain Res 37(2):387-98, 1979
- Imamizu H, Miyauchi S, Tamada T, Sasaki Y, Takino R, Puetz B, Yoshioka T, Kawato M : An internal model of a novel tool: a functional MRI study. Kawato Dynamic Brain Project

Technical Report vol. KDB-TR98, 1998

Imamizu H, Miyauchi S, Tamada T, Sasaki Y, Takino R, Putz B, Yoshioka T, Kawato M : Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature* 403(6766):192-5, 2000

Imamizu H, Shimojo S : The locus of visual-motor learning at the task or manipulator level: implications from intermanual transfer. *J Exp Psychol Hum Percept Perform* 21(4):719-33, 1995

Ito M : Internal model visualized. *Nature* 403(6766): 153-4, 2000

Ito M, Kano M : Long-lasting depression of parallel fiber-Purkinje cell transmission induced by conjunctive stimulation of parallel fibers and climbing fibers in the cerebellar cortex. *Neurosci Lett* 33:253-58, 1982

Ito M, Sakurai M, Tongroach P : Climbing fibre induced depression of both mossy fibre responsiveness and glutamate sensitivity of cerebellar Purkinje cells. *J Physiol (Lond)* 324:113-34, 1982

Kano M, Rexhausen U, Dreessen J, Konnerth A : Synaptic excitation produces a long-lasting rebound potentiation of inhibitory synaptic signals in cerebellar Purkinje cells. *Nature* 356: 601-604, 1992

川人光男 : 小脳外側部の内部モデル, ヒトの知性の計算エンジン : 想像, コミュニケーション, 言語, 思考, 意識, 別冊数理科学, 特集 : 脳科学の最前線-数理モデルを中心として, サイエンス社, 1997, pp194-208

Kawato M : Internal models for motor control and trajectory planning. *Curr Opin Neurobiol* 9(6):718-27, 1999

Kawato M, Furukawa K, Suzuki R : A hierarchical neural-network model for control and learning of voluntary movement. *Biol Cybern* 57(3):169-85, 1987

Kawato M, Gomi H : A computational model of four regions of the cerebellum based on feedback-error learning. *Biol Cybern* 68(2):95-103, 1992

Kawato M, Gomi H : The cerebellum and VOR/OKR learning models. *Trends Neurosci* 15(11):445-53, 1992

Kitazawa S, Kimura T, Yin PB : Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature* 392(6675):494-7, 1998

Kobayashi Y, Kawano K, Takemura A, Inoue Y, Kitama T, Gomi H, Kawato M : Temporal firing patterns of purkinje cells in the cerebellar ventral paraflocculus during ocular following responses in monkeys II. Complex spikes. *J Neurophysiol* 80(2): 832-48, 1998

- Middleton FA, Strick PL : Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science* 266(5184): 458-61, 1994
- Nitschke MF, Kleinschmidt A, Wessel K, Frahm J : Somatotopic motor representation in the human anterior cerebellum. A high-resolution functional MRI study. *Brain* 119(Pt 3):1023-9, 1996
- Raichle ME, Fiez JA, Videen TO, MacLeod AM, Pardo JV, Fox PT, Petersen SE : Practice-related changes in human brain functional anatomy during nonmotor learning. *Cereb Cortex* 4(1):8-26, 1994
- Raymond JL, Lisberger SG, Mauk MD : The cerebellum: a neuronal learning machine? *Science* 272(5265):1126-31, 1996
- Sakurai M : Synaptic modification of parallel fibre-Purkinje cell transmission in in vitro guinea-pig cerebellar slices. *J Physiol (Lond)* 394: 463-80, 1987
- Shidara M, Kawano K, Gomi H, Kawato M : Inverse-dynamics model eye movement control by Purkinje cells in the cerebellum. *Nature* 365(6441): 50-2, 1993
- Strick PL, Hoover JE, Mushiake H : Evidence for "output channels" in the basal ganglia and cerebellum. *Role of the Cerebellum and Basal Ganglia in Voluntary Movement*, Elsevier Science BV, 1993, pp 171-180
- Thach, WT: On the specific role of the cerebellum in motor learning and cognition: Clues from PET activation and lesion studies in man. *Behavioral and Brain Sciences* 19: 411-431, 1996
- Wolpert D, Kawato M : Multiple paried forward and inverse models for motor control. *Neural Networks* 11: 1317-1329, 1998
- 山本憲司, 小林康, 竹村文, 河野憲二, 川人光男 : 垂直方向追従眼球運動の適応シミュレーション : 小脳皮質の計算モデルは運動学習を再現できるか? 電子情報通信学会技術報告 NC97(131):229-236, 1998

到達運動の最適化と誤差の信号

北澤 茂 (電子技術総合研究所、情報科学部)

1. はじめに

ここでは、物に手をのばす、というありふれた運動を話題にする。感動とは無縁にみえて、この運動が意外にも奥深いのである。実は物に手を伸ばすたびに、私達は「最もなめらかな運動」を実現していたのである(Flash & Hogan, 1985; Uno et al., 1989)。宇宙のようなマクロな世界や、素粒子のようなミクロの世界ならいざ知らず、等身大のありふれた動作のなかに、たった一言で表すことのできる原理が隠されていたという事実に、大学院に入ったころの私は深く感動した。こんなきれいな原理が、屋根の上に屋根を重ねてきた神経系の、どこにどうやって入り込むことができたのか？

鍵を握るのは小脳である。実際、小脳を失うと「最もなめらかな運動」もたちどころに失われる。また、乳児の到達運動は初めはジグザクなので、「最もなめらかな運動」は学習によって獲得されるはずだが、小脳には学習にうってつけの構造と可塑性が備えられている。10年前の感動と問に答を求めて、私は小脳で運動学習の研究を続けてきた。

2. 到達運動の美しさ

目標に手を伸ばす到達運動の目的は、許される誤差の範囲内で手を目標の傍に運ぶことだが、目標に達するための運動は数限りなく存在する。実際に実現しているのはどんな運動なのか。実際に手の軌跡を計測すると、始点と終点を結ぶ直線に近い緩やかな曲線である(図1b)。速度曲線はいわゆるベル型である(図1c)。

a. 躍度最小モデル (Flash と Hogan, 1985)

この軌道の持つ意味について、Flash と Hogan は 1985 年に「躍度最小モデル」を提案した。彼らは、実際の手の運動が、手先の位置を時間で3回微分した「躍度」の2乗を運動の開始から終了まで積分した量(評価関数)が最小になるような軌道でよく近似できることを発見した。手先の位置座標を(x, y)とすると、最小にすべき評価関数は

$$C_J = \frac{1}{2} \int_0^{t_f} (\ddot{x})^2 + (\ddot{y})^2 dt$$

と書ける。但し、 t_f は運動時間である。位置を時間で1回微分すると速度、2回で加速度だから、3回微分した「躍度」は加速度の時間変化率、ということになる。手が石ころのような物だとすれば、力は加速度に比例するから、「躍度」は力の時間変化率と比例し、躍度最小モデルから得られる軌道は、力の変化率がなるべく小さくなるような軌道ということになる。

このモデルから得られる軌道は、始点と終点を結ぶ直線で、速度は中点で最大となる 4 次関数である。たったこれだけの単純な原理で、到達運動の主要な性質をほぼ説明する。

b. トルク変化最小モデル (Unoら, 1989)

しかし、手は決して石ではない。腕の先についていて、運動の指令は腕の筋肉に対して送られる。Unoら(1989)は評価関数が腕の力学的性質を反映すべきであると考え、トルク変化最小モデルを提案した。これは、次式のような関節各々に作用するトルク(力のモーメント)の時間変化の 2 乗和の積分を評価関数として選んだ。

$$C_T = \frac{1}{2} \int_0^{t_f} \sum_i (\tau_i)^2 dt$$

但し、 τ_i は i 番目の関節に作用するトルクである。

トルク変化を最小とするような運動は、到達運動の始点、経由点、終点の位置に応じた曲がり具合やバネに抗して行う運動の曲がり具合にいたるまで、実際の運動を驚くほどよく再現する。

c. 終点誤差分散最小モデル (Harris と Wolpert, 1998)

運動を滑らかにすることのメリットは何だろうか。関節への負担が減って怪我をしにくくなるかもしれない。しかし、滑らかに最適化された運動を作り出すには計算時間などの相応のコストを払わなければならない可能性がある。従って、生存競争を勝ち抜くのに有利になるかどうか、直感的には明らかではない。生物学的にもっともらしい目的に基づいて滑らかな運動を説明することができるに越したことはない。

最近 Harris と Wolpert(1998)は到達運動の目的に立ち返るとともに、神経系の情報伝達で避けることのできないノイズの性質を使って、実現されている滑らかな運動に対して見事な説明を与えることに成功した。彼らは「到達運動の目的は終点での誤差を小さくすることにある」とし、さらに神経系の制御信号には平均 0 で制御信号の大きさに比例した標準偏差のランダムなノイズが入ると仮定した。この仮定は運動ニューロンの平均発火頻度と平均発火頻度からのずれとの関係に留まらず、大脳視覚野のニューロンなど神経系一般に観察される実験事実である。時刻 t における腕の状態(位置、速度、加速度など)をまとめて x_t 、時刻 t における制御信号(各筋肉への運動指令)をまとめて u_t 、加わるノイズを w_t と書くと、系の運動方程式を離散時間に直した状態方程式は適当な行列 A , B を用いて、

$$x_{t+1} = Ax_t + B(u_t + w_t)$$

の形に書くことができる。彼らはさらに各時刻のノイズ w_t が互いに独立であると仮定して、終点(時刻 T) からある程度の時間幅(R)の区間の手の位置の誤差の分散(V_t)の和

$$C_T = \sum_{t=T}^{T+R} V_t$$

を最小にするような制御 u_t と実現される運動 x_t を求めた。この終点での誤差の分散を最小化するような運動が、なんと現実の運動、即ち滑らかな運動とよく一致したのだ。直感的に言えば、

大きい信号は大きいノイズを伴い、終点に大きな誤差を生むのでできるだけ避ける。もう少し詳しく言えば、運動開始時の制御信号のノイズは後々まで積分されて大きな誤差を生むから運動始めの指令には、より大きなペナルティーを課す。このようにして大きい信号を避けるということは信号の上昇とその後の下降の幅が小さいということにつながる。つまりは信号の変化率が少くなり、結局のところ「滑らかさ」もかなりの程度実現される。

Harris と Wolpert のモデルによって、到達運動の目的である「手を正確に目標に運ぶ」と「優雅に運ぶ」ことが結びついた。生体の抱える「ノイズ」という負の宿命が「優雅さ」と「正確さ」を結び付けているところがまことに趣深い。もう一つこのモデルが優れている点は、生体は終点の誤差だけを気にしていればよい、ということである。運動全域での滑らかさを確保するために緩和計算や繰り返し計算をする必要はなくなる。このモデルは、終点付近の誤差分散を減らすような学習を続けることができれば、それだけで優雅な運動に近づいていくことを保証する。しかし、このような学習が可能であるかどうか、また脳で実際にどうやって実現しているのか、という難しい問題が解決されたわけではない。以下1つ1つ問題点を洗い出していこう。

3. どこで？

きっと小脳である。少なくとも小脳が深く関わっているだろう。古典的な「小脳症状」としてよく知られているように、小脳に障害があると手は目標を外し(測定障害、dysmetria) 運動は協調性を失う。「正確さ」だけでなく「優雅さ」も失われてしまうのである。これは前節の評価関数を用いて定量的に示すこともできる。例えば、小脳の中位核と歯状核を破壊したネコでは、躍度を用いた評価関数の値が破壊前の10倍にも大きくなる。また、「プリズム順応」と呼ばれる到達運動の再学習過程は小脳の異常で著しく障害されるので、到達運動の学習場所としての必要条件も満たしている。ここでは、小脳で到達運動の終点誤差分散を減らすような学習が行なわれていると仮定することにしよう。では、どうやって？

4. どうやって？

小脳皮質の唯一の出力細胞であるプルキンエ細胞には2系統の興奮性入力が入る(図2a)。一つは苔状線維、顆粒細胞の平行線維を経る入力でプルキンエ細胞に数十から数百ヘルツの頻度の単純スパイクを生じさせる(図2b)。一方、もう1つの入力系、下オリーブ核からの登上線維を経る入力はプルキンエ細胞に平均発火頻度1Hz未滿の複雑スパイクを生じさせる(図2c)。プルキンエ細胞の細胞外から微小電極で電気活動を記録すると、単純スパイクとは全く異なる「複雑な」波形を示すので単純スパイクと区別することができる。

Ito (1970)は30年も前に平行線維からの入力が運動の制御に用いられ、登上線維からの入力は運動の誤差を与えて制御信号の改善に寄与すると推定していた。そして、登上線維経由の入力は平行線維経由の入力と相互作用して、平行線維とプルキンエ細胞の間のシナプス効率を長期にわたって変化させることが学習の実体であると仮定した(Marr, 1969; Ito, 1970; Albus, 1971)。この仮定の下では、登上線維経由の信号が運動の「結果」を与え、「原因」を作った平行線維と

プルキンエ細胞の間のシナプス結合を修正することによって学習が進行するということになる。その後これら2種類の入力の間相互作用に基づくシナプス効率の長期的な変化が実在し、抑制性であること(長期抑圧)が証明された(Itoら, 1982)。この枠組みに則れば、到達運動の制御は単純スパイクが担い、複雑スパイクは到達運動の誤差を表現して学習に寄与するはずである。単純スパイクの発火を上肢運動の制御信号と関連付けることに基本的に異論はない。しかし、複雑スパイクの役割に関しては、誤差を伝えて学習に寄与するという学習説の他にも対立する仮説が提唱されてきた。複雑スパイクの発火頻度が運動の開始時に上昇するというデータから、運動の学習というよりむしろ制御に役立っているのだ、という説である。

a. 複雑スパイクは誤差か行く先か？

到達運動の学習に貢献するならば、登上線維信号は到達運動の誤差、それも運動開始時ではなく運動の終点での誤差を表現してほしい。しかし、過去の研究を調べるほどに、むしろ運動の開始時に行く先に応じて出現して運動の制御をオンラインで補助する可能性も浮かび上がってきた。そこでわれわれは、視覚目標に対して手を伸ばして触れる到達運動をサルに訓練して行わせ、その運動の間に、小脳のプルキンエ細胞の電気活動を記録して、複雑スパイクが運動の行く先を表現するのか、運動の終点の誤差を表現するのか調べることにした(Kitazawaら, 1998)。

サルの目の前にはサルの視覚情報を制御するために液晶シャッターを置いた。サルが目の下20cmに固定されたボタンを指で押えると、液晶シャッターが開き(図3a、試行開始)、予告音に引き続いて目の前方20cmに置いた画面上5x5cmの範囲内のランダムな位置に十字の目標が現われる(目標提示)。サルは目標提示から240 msec以内に指を乗せていたボタンから手を離し(運動開始)手を離してから300 msec以内に画面に触れ(運動終了)なければならない。できるだけ早く反応し、限界に近いスピードで運動しないと間に合わない条件を課した。さらにボタンを離すと同時に眼前の液晶シャッターを閉鎖し、視覚情報なしに到達運動を行なわせた。シャッターは指が画面に触れると同時に再び開き(開)、到達位置を保持したままの手と目標の位置(運動の誤差)が300 msの間視覚情報として与えられた。試行の終了時には、報酬として誤差の量に反比例した量のジュースを与えた。正確に目標を捉えた場合には、たくさんのジュースがもらえるが、はずれてしまうと急激にジュースの量は減る。約30cmの距離を最高約2m/sの速度で行なう厳しい条件の到達運動であるが、終点の誤差の標準偏差は約5 mmと極めて正確な運動が実現された。

図3bにはこの運動の前後のどんなタイミングで単純スパイクや複雑スパイクが出たのか、第V小葉の1個のプルキンエ細胞から記録した20試行分のデータを示した。縦線が単純スパイク、丸が複雑スパイクの出現を示している。この細胞の単純スパイクは運動の終了直前に減って、直後に急増している。ところが、複雑スパイクは、1試行中に1回でるかでないか、出ても2回までである。なんとも頼りなく、きまぐれな出現の仕方である。こんな気紛れな信号に行く先や誤差の情報が含まれているとは一見すると信じがたい。

複雑スパイクはなかなか出ない。これが謎の原因である。ならば、十分な数が出るまで記録し

続けるまでである。図 3b にデータを示した細胞からは 1382 回記録を続けた。この 1382 回の記録の、運動の開始時の 100 ms に注目すると、133 回の試行で複雑スパイクが出現していた。目標はスクリーン上のランダムな場所に出現するので 1382 回の試行でサルが触った点（運動の行く先）は、四角い領域を万遍なく覆っている（図 4a、点）。ところが、運動の開始時に複雑スパイクが出た 133 回の試行に限って丸をつけると、丸は右下に集中していた。丸の数を数えると、左上の第 2 象限には 14 個しかないが、右下の第 4 象限には 56 個あるので、運動の開始時に複雑スパイクが出た場合には、サルがスクリーンの右下に触れる確率が高い、ということになる。従って、運動開始時に複雑スパイクが出たら、多分行く先は右下だろうと予想できる。つまり、運動開始時の複雑スパイクは行く先の情報をもっている。

次に、同じ細胞で、運動の終了直後の 100 ms に注目しよう。この時間内には 92 回の試行で複雑スパイクが生じていた。今度はサルが目標をどれくらいはずしたか、運動の誤差に注意する。1382 回の試行の誤差は、目標のまわりにどの方向にもほとんど均等に分布していた（図 4b、点）。ところが、運動終了直後に複雑スパイクが出現した 92 回の試行に限って見ると、誤差は左上に集中していた（図 4b、丸）。従って、運動の終了時に複雑スパイクが出た場合には、サルが目標を左上にはずしただろう、と想像できる。つまり誤差に関する情報が得られるわけだ。同じ細胞の複雑スパイクが、出る時機に応じて行く先と誤差の両方の情報をそれぞれ伝えていたのである。

この行く先と誤差に関する情報は、ビット単位で定量化することができる。図 5a にこの細胞の複雑スパイクが表現していた行く先（灰色線）と誤差（黒線）に関する情報量の時間変化を示す。運動開始時に行く先の情報が 1 つピークを作り（行く先）、運動終了直前から直後に誤差の情報（黒線）が 2 個のピーク（誤差 1, 2）を作っている。

図 5b には同様の解析を行なった 50 個の細胞の情報量を加算した結果を示した。50 個の細胞の 7 割（34 個）が 3 つのピークの少なくとも 1 つに貢献していた。単純スパイクの発火頻度の変化はそれぞれまったく異なる変化を示したにもかかわらず、「行く先」（灰色線）または「誤差」（黒線）に関する情報量が現われるタイミングは驚く程再現性が高かった。すなわち、「行く先」の情報は運動開始直後から運動前半にかけて 1 個のピークを形成し、一方「誤差」の情報は運動終了直前から直後 250 msec にかけて 2 個のピーク（誤差 1, 2）を形成した。小脳の複雑スパイクは運動の開始時には「行く先」を、終了直前直後には「誤差」を表現していたのである。つまり、誤差も行く先も、というのが私たちの得た解答である。

b. 終点の誤差は時間を遡るか

われわれはこうして、到達運動の終点の誤差が、小脳の登上線維信号で表現されていることを知った。これで「どうやって？」に対する大まかな答えが描けるかもしれない。つまり、登上線維が表現する終点の誤差の情報（複雑スパイク）が平行線維入力と相互作用して、平行線維入力とブルキンエ細胞のシナプス効率に可塑的な変化（長期抑圧）を起し、終点の誤差分散を最小化するような制御信号（単純スパイク）を作り出すことによって、優雅な運動の制御が実現される、という筋書きである。

しかし、こうもうまく学習が進むものか。終点の誤差の情報が、「原因」を作ったシナプスを時間を遡って探し出し、さらには正しい方向（終点の誤差を減らす方向）に矯正しない限り、優雅な運動に近づくことはできないだろう。少なくとも、遅れて戻ってくる登上線維信号が、先行して入ってきた平行線維入力とプルキンエ細胞の間のシナプス効率を変化させることができなければならない。

小脳スライス標本を用いた Chen と Thompson(1992)によると、登上線維入力は平行線維の入力より 250 ms 遅れても長期抑圧を生じさせるという。さらに興味深いことには、125 ms の遅れでは 250 ms の遅れの場合の約 3 分の 1 の抑圧しか生じず、同時刺激 (0 ms) と登上線維の 250 ms 先行刺激または 750ms 遅れた刺激もほとんど長期抑圧を生じなかった。彼らの得た結果は、250ms 程度遅れて戻る登上線維信号がむしろよく長期抑圧を起こすことを示唆している。

では、運動の誤差を伝える登上線維信号は、実際の運動からどれくらい遅れて戻ってくるのだろうか。もう一度図 5 を見る。誤差の情報量には 2 つの成分 (図 5b) があつた。先行する成分 (誤差 1) は、視覚入力遮断されている運動の終了前から運動終了直後にかけて、2 つ目の成分 (誤差 2) は運動終了後約 100 ms 後から 300ms 後にかけて出現した。シャッターを開いて誤差を見せるタイミングを遅らせると 2 つ目の成分は視覚入力の遅れの分だけ移動した。従って、この成分 (誤差 2) は視覚性に検出された誤差の成分と考えられる。立ち上がり潜時は 100 ms、ピークまでの潜時は 180 ms であつた。一方、先行する誤差成分 (誤差 1) は視覚由来ではない。これは体性感覚と運動司令の遠心性コピーを入力として予測的に作られる誤差の情報だろうと考えている。

Chen and Thompson(1992)に基づいて、これらの誤差成分を時間軸上で 250 ms 溯らせてみよう。極めて大雑把ではあるが、これら 2 つの誤差情報成分は全体として十分に運動制御の期間を覆いうる。しかし、視覚性の (誤差 2) 成分は主に運動の後半を覆うのみである。従って、到達運動の終点で与えた視覚性の誤差情報だけでは運動の前半を制御するシナプスの効率を効果的に変化させることはできないだろう。つまり、運動の前半の制御を向上させるためには予測性の誤差成分 (誤差 1) が極めて重要な役割を果たしているはずだ。

予測性の誤差成分 (誤差 1) はどこから来るのか。小脳自身が有力な候補の 1 つである。実際、まだ数は少ないが、到達運動の途中で終点の誤差の情報を持った単純スパイク (複雑スパイクではないことに注意) を出すプルキンエ細胞を見出している。誤差を予測するプルキンエ細胞の単純スパイクの出力が歯状核、小細胞性赤核、下オリーブ核を経て登上線維信号として再び小脳に入力して、運動の制御に関わるプルキンエ細胞の学習に使われている可能性もある。

シナプスレベルの遡り能力に加えて、誤差を予測する機能を使えば、到達運動終点の誤差の情報を運動制御の時間全体に遡って行き渡らせることができるだろう。

5. 到達運動に到達できるか？

ごくありふれた到達運動は、滑らかさにおいてほぼ「最適化」された運動だった (Flash と Hogan, 1985; Uno ら, 1989)。その「最適化」は終点の誤差を小さくしようとする生物学的に合

理的な目的に適う、あるいはむしろ終点の誤差を小さくしようと脳が努めた結果として「最適化」がもたらされた、と考えられることが示された(Harris と Wolpert, 1998)。そして、最適化された運動制御と学習による獲得の鍵を握る小脳で、運動中には極端に寡黙な登上線維信号に確かに終点の誤差の情報が含まれていること、しかも予測性の成分も含まれていることが明らかにされた(Kitazawa ら, 1998)。あとはこの登上線維信号を使って誤差を減らすような学習をするだけである。とはいえ、本当に終点の誤差(分散)が減る方向へ、終には最小化する運動制御を実現するように学習は進んでいくのだろうか。また、その学習は小脳だけで行ないけるのだろうか。

小脳に運動の「目標軌道」が入力し、登上線維信号として運動の誤差を補正するフィードバック誤差信号が入力するならば、小脳は目標軌道を実現するような運動制御信号を出力するように学習することは、極めて明快に理論化されており(フィードバック誤差学習) また現実のロボット制御でも華麗な成功を収めている(川人, 1996)。この過程で小脳には目標軌道を運動出力に変換する「逆モデル」が獲得される。

では、到達運動において小脳に入力する「目標軌道」は何か。「終点誤差分散最小軌道」が実現されている以上、小脳に正確な「逆モデル」があるならば、逆モデルへ入力するのは「終点誤差分散最小軌道」そのものとなる。ということは、すでに脳のどこかが「終点誤差分散最小軌道」を知っていることになる。では「終点誤差分散最小軌道」を獲得するのはどこか。軌道計画と制御をロボット制御のようにシリアルに行なうと仮定する限り、阪口らが指摘するように賢い軌道計画を行なう場所と仕組みを探す必要が生じる。

Wada と Kawato(1993)は軌道計画と制御を同時に行い、トルク最小軌道を生成する目的で逆モデルと順モデル(制御信号を腕の軌道に変換するモデル)をループ状につないだ巡回型の神経回路(図6)を提案した。実際の制御に先行してこのループを回して運動制御信号の最適化を図った上で制御信号を出力する、というのが元々のアイデアである。ここでは運動開始前にループを回す最適化計算を放棄してみる。ループを回し始めると同時に制御を開始するのである。この場合には、もはや運動制御に先行する軌道計画は存在せず、ループ全体として「誤差分散最小制御」を獲得できればよいことになる。

さらに「逆モデル」に目標の位置が運動の制御期間に亘って定常的に入力し、この目標の入力と同時に制御が始まると仮定する。もはや「目標軌道を運動出力に変換する」という厳密な定義からははずれるので、「逆モデル」を制御ユニットと呼びかえる(図6)。こうしてできた図6を素朴に読むと、1)制御ユニットに目標位置が入力して制御開始、2)腕の初期状態と目標位置を元に始めの運動制御信号を出力、3)運動制御信号を入力として順モデルによって腕の状態を更新、4)更新された腕の状態と不変な目標位置を使って次の制御信号を出力、という具合に巡回型のネットワークで運動制御のための信号時系列が生成されていく。そしてこの制御ユニットが登上線維信号に基づいて学習を進めるのである。

図6のモデルでは登上線維信号に「誤差」と書かず「目標と手の差」が入力するとしたことにも注意されたい。終点近くでは目標と手の差は誤差そのものである。一方、運動開始時の「目標と手の差」は手から「行く先」へ向かうベクトルである。つまり、われわれが得た行く先の情報

と誤差の情報は「目標と手の位置の差」として統一的に理解することができる。図4の細胞では行く先の適方向は右下(図4a)で、誤差の適方向は左上(図4b)とほぼ逆を向いていた。しかし、手から目標へのベクトル誤差として見直すと、いずれも右下を向いて方向が一致する。この関係は行く先と誤差の情報両方を持つ他の細胞でも基本的に満足されていたので、登上線維信号が終始一貫して手と目標位置の差を知らせ続けていると考えうる。

知りたいのはこの巡回型のネットワークの中に、ベクトル誤差を伝える登上線維信号だけに頼って終点の誤差分散を最小にするような制御を自動的に埋め込むことができるか、という問題に対する答えである。直感的にはとても筋が悪い気がする。誤差の平均が0になる方向へ降りていくとしても、分散が小さくなって、しかも最小に収束するなどということは考えがたい。でもでも、「収束」をあきらめてしまえば道が開けるのではないか。終点の誤差の平均が0であるような無数のシステムの中で考えると、ランダムノイズのおかげで生み出される終点の誤差に応じてシステムはブラウン運動するだろう。となると、終点の誤差分散できっと拡散係数のようなものが定義できて、拡散係数の小さいシステム(終点誤差分散の小さいシステム)付近の滞在時間が長くなる、なんていうことは言えるのではないか?ノイズのおかげで必ず local minimum から脱出することができて、しかも好ましい制御のあたりで動きにくくなるという、天然の simulated annealing が実現されている、などというのは実験屋の妄想か?いかがでしょう、理論の方々。

文献

- 1) Albus, J.S.: *Math. Biosci.*, **10**, 25-61, 1971.
- 2) Chen, C. and Thompson, R.F.: *Learn. Memory*, **2**, 185-198, 1995.
- 3) Flash, T. and Hogan, N.: *J. Neurosci.*, **5**, 1688-1703, 1985.
- 4) Harris C.M. and Wolpert, D.M.: *Nature*, **394**, 780-784, 1998.
- 5) Ito, M.: *Int. J. Neurol.*, **7**, 162-176, 1970.
- 6) 川人光男： 脳の計算理論， 産業図書， 東京， 1996.
- 7) Kitazawa, S., Kimura, T. and Yin, P.B.: *Nature*, **392**, 494-497, 1998.
- 8) Marr, D.: *J. Physiol.*, **202**, 437-470, 1969.
- 9) Uno, Y., Kawato, M. and Suzuki, R.: *Biol Cybern*, **61**, 89-101, 1989.
- 10) Wada, Y. and Kawato, M.: *Neural Netw.*, **6**, 919-932, 1993.

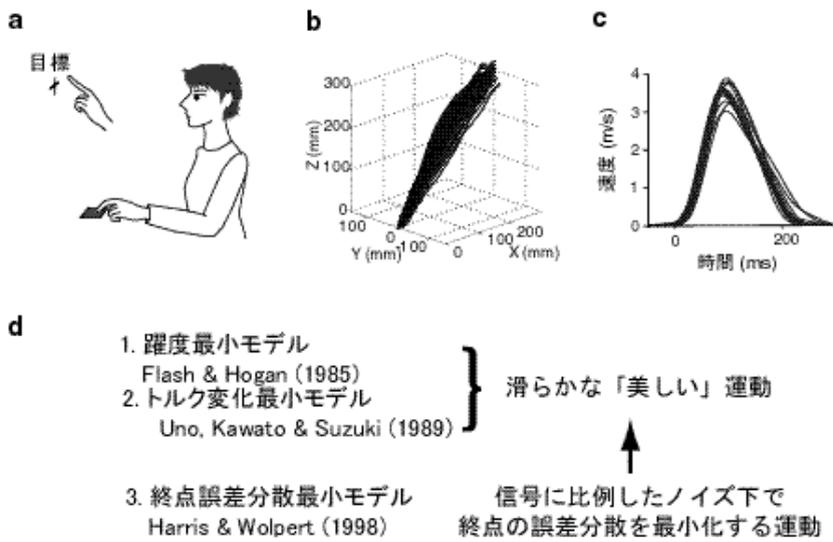


図 1 到達運動とそのモデル

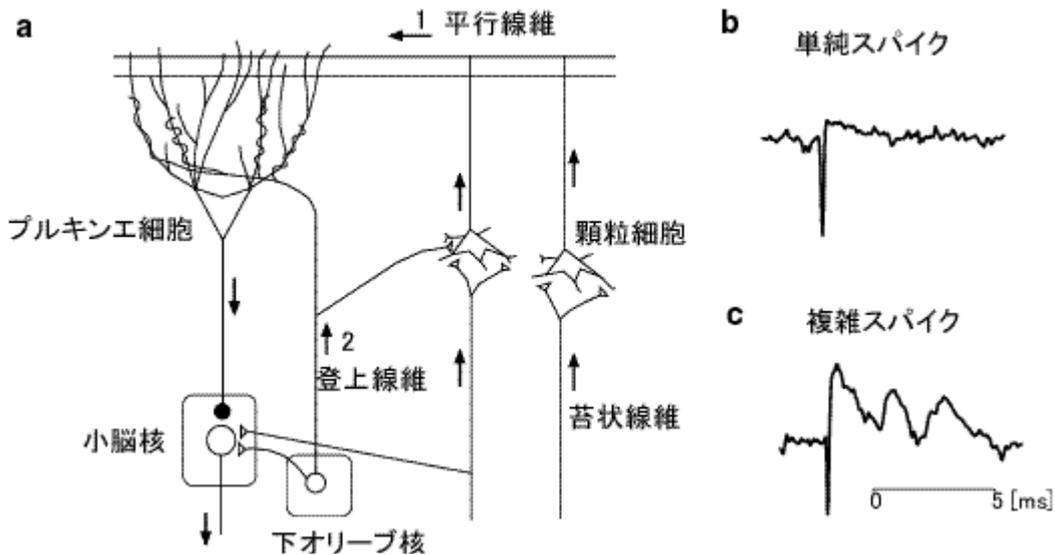


図 2 小脳プルキンエ細胞への 2 種類の興奮性入力とスパイク波形

a : プルキンエ細胞を中心とする小脳の神経回路。興奮性入力に限って簡略に示した (Optican, 1998 改変)。平行線維(1)からの入力は単純スパイク(b)を、登上線維(2)からの入力は複雑スパイク(c)を発生させる。

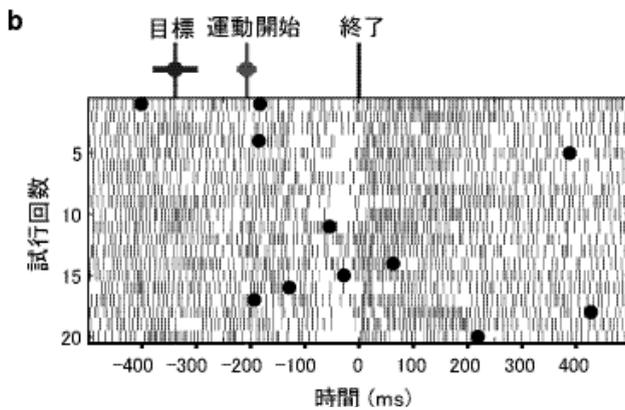
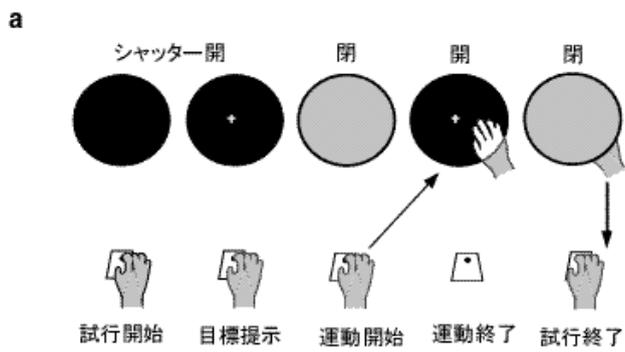


図 3 到達運動中のサルの小脳の電気活動

a : 到達運動課題。 b : 第 V 小葉の 1 個のプルキンエ細胞から記録した単純スパイク(縦線)と複雑スパイク(黒丸)。20 試行分。

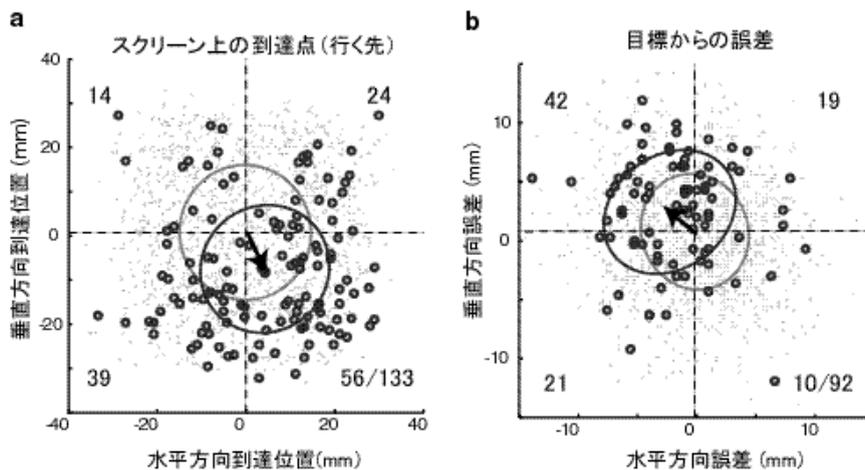


図 4 到達運動の行く先と誤差の分布 (1382 試行分)

1 個のプルキンエ細胞(図 1.4b と同じ細胞)から記録を行なった 1382 回の試行の行く先(a; 点)と誤差 (b)のうち、運動開始時に複雑スパイクが生じた試行(133 試行)の行く先 (a)と、運動終了直後に複雑スパイクが生じた試行(92 試行)の誤差(b)を黒丸で囲んだ。黒丸の分布の偏りに注目。

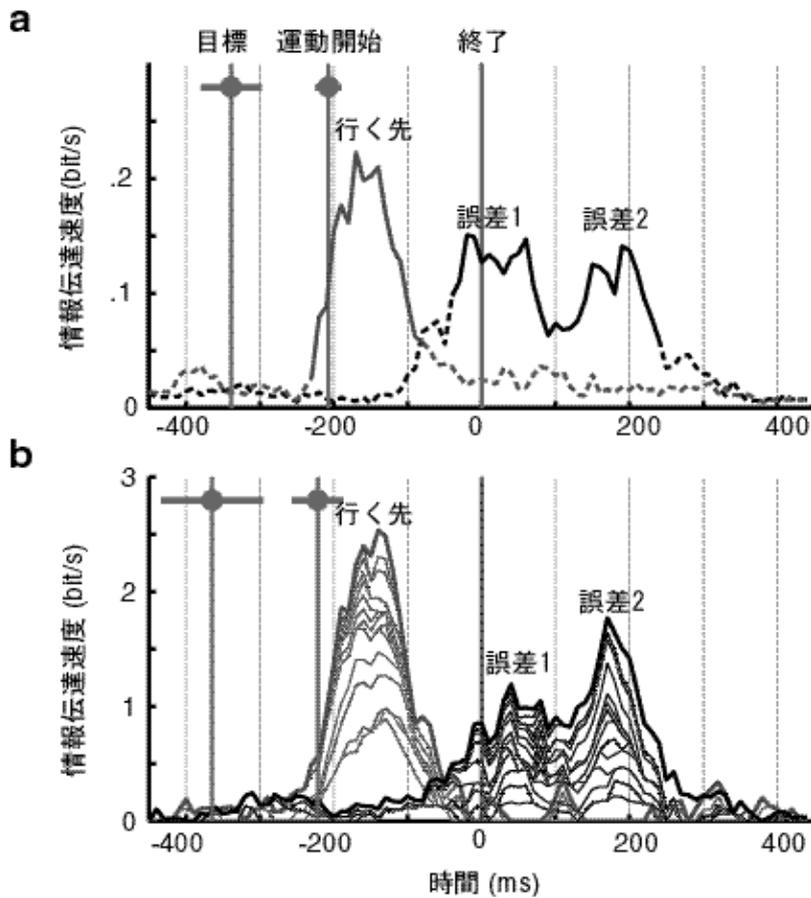


図5 複雑スパイクが表現する行く先と誤差の情報

a:1個のプルキンエ細胞(図1.5と同じ細胞)の表現する情報の時間変化。b:50個の細胞の情報の和。運動開始時に行く先(灰色線)、終了時に誤差(黒線)の情報がピークを作る。平均発火頻度ではないことに注意。

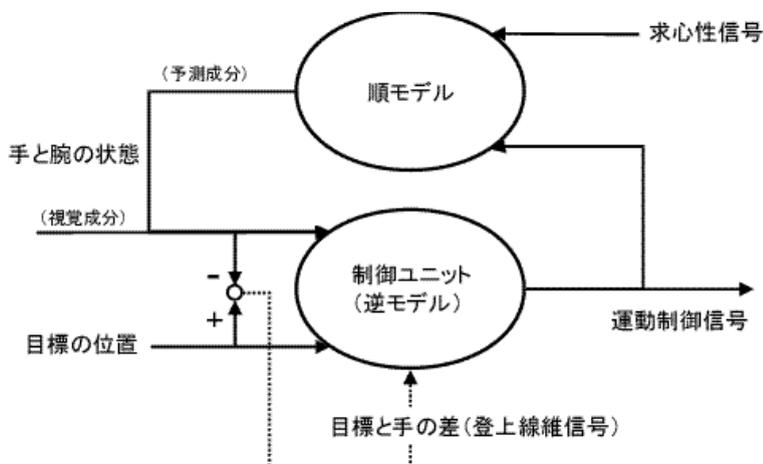


図6 美しい到達運動を獲得するための神経回路モデル

運動制御における高次の問題

— 到達運動と視覚運動変換を例にとって —

阪口 豊

電気通信大学大学院 情報システム学研究科
科学技術振興事業団 CREST

1 はじめに

前節までの議論からわかるように、小脳における内部モデル、情報表現、学習に関する研究はここ 10 年のあいだに大きく進んだ。ここでは、少し視点を変えて、これまでの研究では明確な形で議論されていない問題をいくつか取り上げてみたい。散逸的な内容の文章をテキストに載せるのはどうかとも思ったが、このスクールは「研究になるかならないのかよくわからない問題」を議論するのにちょうどよい機会でもあるので、あえて原稿としてまとめてみた。興味のもった方と共同研究をはじめるときかけになれば、筆者にとってはこの上ないことである。なお、筆者の勉強不足のために内容に誤りのある部分やすでに同種の研究が行なわれている部分があるかもしれないが、その場合はご指摘いただきたい。

2 到達運動のプリズム適応メカニズム

小脳における運動学習のメカニズムは、通常「教師あり学習」という枠組みで捉えられる。この考え方は、小脳パーセプトロン説からフィードバック誤差学習モデルまでの流れの中で一貫したものであり、今日誰もが認める小脳の姿であるといえよう。さて、小脳を「教師あり学習モジュール」とみなして運動系のモデルを組み立てようとしたとき、必ず明確にしないといけない問題として「小脳は何を学習しているのか」、「教師信号（誤差信号）はどのようにして与えられるのか」の二つが挙げられる。

この点で、OFR(ocular-following reflex)の適応モデルは美しく、また明快である。すなわち、小脳は、フィードバック制御回路から誤算信号を受け取ることにより眼球運動系の逆ダイナミクスを学習する（川人さんの節を参照）。それでは、上肢の到達運動における小脳の役割はどうであろうか？以下では、視覚的に与えられた目標に向けて手を動かす視覚誘導性到達運動を例にとって考える。

プリズム眼鏡などにより視覚環境が変形された状態で到達運動を行なうと、最初は正しく目標に到達することができないが、やがて正しく運動できるようになる（プリズム適応）。小脳に疾患のある患者ではプリズム適応が生じないことから、一般にプリズム適応には小脳が必要であると考えられている。さらに、小脳の複雑スパイク（教師信号に相当）に到達運動の誤差情報が含まれていること（北澤さんの節を参照）、マウス操作の学習において小脳に学習を反映する活動が現れること（今水さんの節を参照）から、視覚運動変換の適応や学習に小脳が関与していることはまず間違いがない。

しかし、以下で議論するように、到達運動の適応メカニズムは、OFRの適応メカニズムと比べるとはるかに複雑でわからない部分が多い。これは、上にあげた二つの問題がいずれも解決されていないからである。

2.1 小脳の部位と適応

この問題を議論するにあたり、まず、小脳が複数の部分から構成されていることを意識しておく必要がある。小脳は大きく、1) 脳幹や前庭器官と連絡して眼球運動や定位運動に関与している内側部（前庭小脳）、2) 脊髄や赤核などと連絡して身体の運動に係している中間部（脊髄小脳）、3) 大脳皮質と連絡し

て運動計画などに関与している外側部の三つに分けられる [1]。これらはその出力先も異なっている（順番に、室頂核、中位核、歯状核）ことから、それぞれ異なる役割を担っていると考えられている。ちなみに、OFRの研究では ventral paraflocculus と呼ばれる前庭小脳の一部、今水さんの研究では外側部、北澤さんの研究では中間部と外側部にまたがる領域においてそれぞれ神経活動を計測している。

脊髄小脳と前庭小脳は運動・感覚系と密接に連絡していることから、筋骨格の性質を反映したモデル（例えば、順・逆ダイナミクスモデル）を保持する場所として有力であろう。OFRのモデルは前庭小脳を対象としたものであり [2]、また、脊髄小脳に着目した到達運動のモデルも提案されている [3, 4]。一方、大脳皮質と相互に結合していること外側部は、より抽象度の高い機能に関与していると考えられる。例えば、マウスを自由自在に操るには自分の身体以外の因果関係を獲得する必要があるため、このような機能の獲得に小脳外側部がかかわっていることはうなづける。小脳が認知機能にも関わっているという近年の主張は、外側部がもつこのような高次の役割に着目したものであるといえる。

このような準備の下で、到達運動のプリズム適応を担っているのが小脳のどの部分であるかを考えてみよう。これにはいろいろな可能性があるが、おそらく小脳の異なる部位は視覚・運動変換の異なる段階に関与しているのではないだろうか。例えば、到達運動に関してキネマティクスとダイナミクスの学習が独立していることを示す心理実験 [5] が報告されているが、このような知見も上の考え方と整合する。もう少し具体的にいくつか可能性について考えてみよう。脊髄小脳は身体の内部モデルと関連が深いので関与している可能性が高いが、仮にダイナミクスモデルを保持しているだけならば、プリズム適応には（ダイナミクス自体は変化しないので）関与していないこともありうる。逆に、前庭小脳は、手先の運動に直接関与しないものの、到達運動の目標情報を得る上で眼位が重要なことや手の動きが姿勢制御と切り離せないことを考えると、間接的に関与している可能性がある。運動計画と密接に関係している外側部は、まず間違いなく関与しているであろう。このように、一口に「小脳が適応に関与している」といっても、その関与の仕方にはさまざまな可能性があり、プリズム適応における小脳の役割を理解するにはそれらを一つ一つ吟味しなければならない。

このような作業を実際に進めるには、到達運動の制御がどのような段階から構成されていて、小脳がどこに位置づけられるのかを示す計算モデルや仮説が必要である。以下では、このような計算モデルの構成にかかわる考察と実験の試みについて述べていきたい。

2.2 誤差信号の由来

次に、誤差信号はどのようにしてもたらされるのかについて考えよう。これは、小脳に対して教師信号を与えている下オリブ核に、どのようにして必要な情報がもたらされるのかという問題でもある。

当然のことではあるが、何が教師信号になるかは何を学習するかに依存して決まる。例えば、プリズム適応において小脳が「関節座標系から視覚座標系への順キネマティクス」を学習していると仮定しよう。すると、誤差信号は「自分の手先がどこに見えるかの予想値（つまり順キネマティクスモデルの出力）と実際に見える手先の位置との差」として与えられる。そのような情報を与える一つの候補は「視覚で捉えた到達点と目標のずれ」であるが（ここでは、目標位置が順モデルの出力で、到達点が正解であることに注意が必要である。正しく運動することに興味が向いてしまうと、手先位置が出力で目標位置が正解であると思ってしまう）、そのほかにも候補はある。運動前野には「手の位置に応じて視覚的な受容野の位置が変化する細胞」が存在するが [6]、このような細胞の出力が誤差情報を提供する役割を担っているかもしれない。このほか、到達運動では通常目標を注視することから、自分の手先が視野のどこに見えるか、という情報も誤差信号の源になりうる。

このように、必要な誤差信号がどのような経路で与えられるのかはそれほど自明な問題ではない。この問題を解き明かす上でも、計算モデルや心理学的な手法の役割は重要である。すなわち、計算モデルの手法を用いてさまざまな構造のモデルを仮想的に設定し、どのような情報が教師信号となりうるのかを検討したうえで、教師信号となりうる情報を統制して適応の進み方に生じる変化を調べる心理実験を行なうのが有効であろう。これらのスクリーニングを行なったうえで神経生理学的な実験を行なえば、適応メカニズムの中身を効率的に調べていくことができる。

2.3 教師あり学習と強化学習

次に、視点を小脳からさらに大きな範囲に広げて、「適応は教師あり学習だけで（つまり小脳だけで）実現されているのか」という問題を考えてみよう。

到達運動の実験やプリズム適応の実験に参加してみるとわかるが、自分の手先が思うように目標に届かないと、人間はどうやれば目標に正しく到達するかと考えて身体を動かすようになる。心理実験ではこのような恣意的な因子をできるだけ小さくするために、反応時間や運動時間に制約を課したり被験者への指示を工夫したりするが、「できるだけ正確に目標に到達したい」という意識が働く以上、試行錯誤的な要素を完全に排除することは不可能である。

「正しく目標に到達したかどうか」という結果の評価には、到達点と手先のずれという「誤差の評価」に加えて、うまくいったかいかかなかったかという「報酬としての評価」も含まれる。ここで、「試行錯誤」と「報酬」という言葉の組み合わせから強化学習を思い出す方も多いただろう。さらに、教師あり学習 - 小脳、強化学習 - 大脳基底核という枠組み [7] を考えると、小脳よりも大脳基底核の働きを中心に考えた方がよいのではないか、という気がしてこないでもない。

しかし、小脳疾患のある患者においてプリズム適応が生じないことを思い出せば、小脳が適応に関わっていることは間違いがない。これらを考えあわせると、「プリズム適応には教師あり学習による成分と強化学習による成分がある」、さらには、「小脳と大脳基底核が協調的に機能することで適応が実現される」と捉えるのが自然ではなかろうか。

以下では、これら二つの学習がプリズム適応に関与しているという仮定の下で議論を進める。それでは、二つの学習はどのような関係にあるのであろうか？また、二つを分離して取り出すことは可能であろうか？

筆者は、小脳での教師あり学習はある決まった入出力関係を学習するために機能するのに対し、大脳基底核での強化学習は文脈における最適なセットを選択することを学習すると考えている（大脳基底核が文脈情報提供の役割を果たすという考え方は新しいものではなく、Thachら [8] もこのような考え方をとっている）。プリズム適応のような視覚運動変換の実験を何度も繰り返すと、被験者は、通常的环境と変形された環境との運動を随意的に切り替えられるようになる。このことから、脳内には複数の運動セットが並列的に形成されることは間違いがないだろう。ここで問題になるのは、だれがそのセットを切り替えるのかということである。筆者の考えは、この切り替えの学習に強化学習が用いられているということである。

この問題は、モジュール型ネットワークの切替え問題の一例であると捉えられる。これまでに提案された有力モデルには、切り替えエキスパートを設ける方法 [9] と順モデル出力と結果の差に基づいて切り替える方法 [10] があるが、ここでは、大脳基底核が切り替えエキスパートの役割を果たしているというモデルを考えたい。小脳と大脳基底核の双方と連絡のある運動前野にムシモルを注入すると適応が起きなくなるという現象 [11] は、このような考え方で説明できるかもしれない。

さらに、この問題に関連して「教師あり学習は自動的に進行するが、強化学習は報酬を意識しないと進行しない」という仮説についても考えてみたい。反射運動である OFR の学習は被験者の意識とは無関係に進行することから、小脳における学習は被験者の意識とは無関係に進行するといえるだろう。これに対し、高次の強化学習は、被験者が報酬を明確に意識しなければ進行しないと思われる。プリズム実験で獲得した複数の環境は随意的に切り替えることができるが、成長して身体が大きくなったとき成長以前の環境での運動を思い出すことは不可能であることも、その違いを示す一例ではないだろうか。つまり、随意的な切り替えができる場合は、内部モデルの学習と内部モデルの切り替えの学習がともに起こり、切り替えができない場合は内部モデルの学習だけが起こるということである。

このように、意識と学習のあいだに関係があるとすれば、被験者に変化を意識させない状況を作り出すことで、強化学習が働かない（つまり教師あり学習だけが働く）状況を作り出し、それにより教師あり学習だけを取り出すことが可能になると考えられる。そのような実験の試みについては次節で紹介する。

この節の要点をまとめると、「適応は小脳だけで実現されている」のではなく、「小脳を含めた複合系によって実現されている」ということである。ここで述べた教師あり学習と強化学習の複合モデルの狙いは、銅谷さん [7] が提唱している小脳、大脳基底核、大脳皮質の学習上の役割分担の構図を土台にして、Thachら [8] のモデルに見られるような小脳と基底核の計算上の役割分担を一緒に考えることにより、両者のダイナミックな関わりを議論することにある。

2.4 変換の知覚が適応に与える影響—心理実験

前節で述べた問題意識の下で、筆者らはプリズム適応の心理実験を始めたところである。ここでは、「プリズム偏位の存在に気づかない状況と気づく状況での適応の違い」を調べる試みを紹介する。また、予備実験の域を出るものではないが、詳しい内容に興味のある方は信学技報 [12] をごらんいただきたい。

この実験では、被験者は目標が提示され次第、できるだけ正確に目標まで手先を動かす運動を行なう。運動は水平面内で行なわれる。恣意的な動作をできるだけ排除するため、反応時間、運動時間には上限が設けられ、それを越えた試行はすべてキャンセルされる。

被験者の眼前には鏡がおかれているため、被験者は自分の手の動きを直接見るができないが、手先位置に応じて移動するレーザスポット光の投影像を鏡を通して観察することができる。ただし、手先位置は運動終了時点でのみ提示され、運動中は表示されない。ここで、スポット光の位置を人為的に操作することで、プリズム偏位と同様の環境を実現する（この実験では左右方向の平行移動のみ）。

被験者が偏位の存在に気づく条件と気づかない条件を作るため、この実験では、偏位量が不連続に増加するステップ型変換と徐々に増加するランプ型変換の二つを設定した。そして、偏位を加えた試行を数十回行なった後、スポット光なしで運動を行なったときの誤差の水平成分を測ることで適応の進み具合を評価した。また、実験終了後に実験中に感じたことを自由に報告してもらい、偏位の気づいたかどうかを調べた。

その結果、筆者らの事前の予想と反し、変換の種類を問わず、偏位の存在に気づく被験者と気づかない被験者がいることがわかった。さらに、変換の種類によらず、偏位の存在に気づいた被験者では適応が生じず、気づかない被験者では適応が生じることがわかった。被験者が偏位の存在に気づくか否かで適応の有無が決まったことは、偏位に気づいた被験者が二つの環境での運動を意図的に切り替えていたことを示唆している。

このほか、誤差の大きさを視覚情報ではなくシンボル情報（数値）として与えても適応が進行することがわかっている。この結果は、感覚系が直接参照できないような形で誤差情報を与えても適応が起こりうることを示しているが、このようにして起きた適応が、通常のプリズム適応と同質のものであるかどうかは議論の余地がある。ただ、先に述べたように、教師あり学習が自動的に生じるのであれば、仮に意識的に修正して行なった運動であっても、そのときの身体の動きから教師あり学習が無意識下で進行すると考えれば、両者が同質である可能性はある。いずれにせよ、相互転移の有無を調べることで、両者の性質を比較することが必要である。

3 軌道計画のメカニズム

ここで、話題を適応から軌道計画の問題に移そう。軌道計画は、到達運動制御に関わる重要な問題の一つである。

到達運動に関する計算理論では、目標軌道は陽に計算されているという考え方が有力である。軌道計画の問題に関しては、脳は運動の終点位置だけを指定しているという終点仮説が Bizzi の実験 [13] により否定され、その後、平衡点軌道（仮想軌道）仮説と逆ダイナミクス仮説のあいだで争いが続いてきたが、これまでに得られている実験結果を公平に判断すれば仮想軌道仮説は極めて不利である。ただ、逆ダイナミクス仮説に関しても、目標軌道が陽に計算されているということを裏付ける実験結果はないように思える（Bizzi の実験結果はフィードフォワード制御の考え方を支持するが、目標軌道の存在を必ずしも支持するものではない）。

目標軌道が陽に計算されるという計算理論の枠組みは、運動制御の過程を、座標変換、軌道計画、運動制御という三つのステップに分けた点と関連している。このような分類はあくまで理論的、概念的なものであり、計算理論は、脳がこれらの処理を独立した形で行なっているとは主張していない。しかし、躍度最小化、トルク変化最小化といった軌道計画規範や、目標軌道が外部から与えられることを仮定した制御系モデルが一人歩きしてしまい、両者の分離が議論の前提になってしまうのは危険である。

OFR に対する計算モデルと同様のレベルで、到達運動のモデルを構築するためには、このような概念上の区別よりも脳の部位と情報処理過程の対応付けを重視したモデルの方が望ましい。それは言い換えれば、

軌道がいつどこでどのようにして決定されているかを議論しなければならないということである。

ここでは、とりあえず「どこで」の問題には脇において、「いつ」「どのように」の二つの問題について少し議論してみたい。なお、到達運動における、小脳や基底核、運動前野の役割を具体的に議論したモデルとして、Houkら [14] のモデルは興味深い。

3.1 動的な軌道計画モデル

「いつ」の問題について議論するために、まず、到達運動の目標を運動開始後に変更したときの被験者の振舞いを紹介したい。到達運動では通常目標を注視するサッカーボールが生じるが、サッカーボール中の視覚能力が低下することを利用すると、被験者に悟られずに目標位置を変更することができる。手先が見えない環境で運動開始後に目標位置が変更された場合、手先が到達するのは変更前の目標であろうか、変更後の目標であろうか？

仮に、軌道計画が運動開始前に終了していても、かつ、到達運動が完全にフィードフォワード制御で行なわれていれば、被験者の手先は変更前の目標に到達するであろう。しかし、実験の結果、手先は変更後の目標に到達することが明らかになった。このことから、Prablancら [15, 16] は、サッカーボール後に得られた目標情報を運動制御に反映させるフィードバック回路が脳内にあると述べている。このフィードバック回路の実体としては、体性感覚フィードバックが一つの可能性として考えられるが、感覚神経を切除した患者においても同じ結果が出たことから [17]、この考え方は成立しない。つまり、視覚フィードバック、感覚フィードバックのどちらも効かない場合でも、手先は新しい目標に到達するのである。

以上の事実から、目標が運動開始後に更新されたときの運動指令の修正は、フィードバック制御ではなく、フィードフォワード制御系に入力される軌道情報の変更によって実現されていると考えるのが自然であろう。ここで、1) 軌道計画処理を運動開始前に終了させてしまい、脳のリソースを軌道計画処理から解放してしまうと、軌道を修正するための処理を新たに立ち上げるオーバーヘッドが大きくなること、2) 軌道修正は、目標が途中で変更されるといった特別な事情がなくても、サッカーボール後の視覚情報を使って指令を微調整する処理は常時行なわれているはずであることを考えると、軌道計画は運動を遂行しながら動的に進められている可能性が高い。

感覚フィードバックに頼らずに軌道を修正するには、内的なフィードバック回路が必要であり、そのためには筋骨格系の順モデルが必要である。また、一連の運動指令パターンを発したときに、それがもたらす軌道を予測するマクロな意味での順モデルも必要であろう。ここでは、そのような二つの順モデル機構を利用して、将来の軌道を予測しつつ適切な運動指令を動的に選択する軌道計画アルゴリズムを考えてみたい。これは、一種の shooting であり、適当な運動指令を選んでそれを生成したときの将来軌道を予測し、その結果を目標位置と比較して運動指令を修正するというものである（トルク変化最小軌道を数値計算で使われる一般化ニュートン法も一種の shooting である）。軌道計画の初期段階である程度の絞り込みが済んでいて、かつ、運動指令の修正範囲に制約があることを仮定すれば（あくまで希望的観測ではあるが）探索問題の負荷はそれほど大きくなりませんと考えられる。

筆者らの研究室では、このような動的な軌道計画モデルの可能性を探っているところである。ここでは一つの案として、確率モデルを用いたアルゴリズムについて述べてみたい。まず、基本的な運動指令パターンの組があらかじめいくつか用意されていると仮定する。各指令パターンに対応して将来の予測軌道が計算されるが、そこには将来の不確実性が含まれており、不確実性の程度が確率モデルとして表されている。目標の情報が与えられると、当初の運動指令パターンが選択される（あるいは複数パターンの荷重和として表しておく）。選択されたパターンをもとに運動指令が出力されると、それを受け取った順ダイナミクスモデルがオンラインで軌道を計算し、それにより軌道が確定するとともに将来軌道の不確実性が減少する。軌道計画モジュールは、各時点で予測される到達点の分布を目標位置と比較し、目標への到達確率が高くなるように運動指令を微調整する。微調整の方法はまだ明確に定めていないが、例えば、複数の運動指令パターンの荷重を動的に調整することがあげられる。これにより、運動途中で目標が変更されても、荷重を変更することにより軌道を動的に変更することが可能である。

「どこ」の問題に関しては、順モデルに基づく腕位置の推定は小脳中間部、将来軌道を予測しつつ運動

指令を絞り込むのは小脳外側部，運動前野，頭頂連合野のネットワーク，文脈に応じて運動指令パターンのセットを切り替えるのが大脳基底核ということになる。

ここでの狙いは，軌道計画と運動制御を一体化したモデルを考えることで，動的な軌道計画修正を可能にし，かつ，Houkらのモデル [14] と同様のレベルで議論することにある．上記のアルゴリズムの是非について意見をいただければ幸いである．

3.2 2重課題パラダイムを用いたアプローチ

到達運動中の脳内過程を探るため，筆者らはこのほかにも次のような実験を行なっている．これは，到達運動中に別の認知課題（音高弁別課題）を実行し，その課題の成績（反応時間）が運動のフェーズごとにとどのように変化するかを調べるものである．

二つの弁別課題をほぼ同時に行なうと，二つの刺激の間隔が短くなるにつれて2番目の刺激に対する反応時間が長くなることから，脳内の弁別遂行過程にはボトルネックが存在することが指摘されている [18]．これに対し，到達運動課題と弁別課題を同時に行なった場合は，運動開始時と終了時に弁別課題の反応時間が若干長くなることが報告されている [19]．このような実験は，到達運動制御の過程が運動系内で閉じたものではなく，他の認知系と干渉する要素をもっていることを示している点で興味深い，これまでの研究の多くは運動制御と「注意」（ここでは脳中枢の情報処理資源という意味）の関係を問うもので，運動制御におけるオンライン計算メカニズムを調べることを目的としたものは少ない [20]．

ここでは，あらためてこのような実験を行なうことにより，運動中の脳内過程を探るための新しい手がかりが得られないかと考えた．従来の実験が試行数を抑えるため，運動の特定のフェーズでのみ音刺激を加えていたのに対し，ここでは音刺激を与える時刻をランダムに決めて，目標提示から運動終了までの過程で音弁別課題の反応時間がどのように変化するかを調べた．予備実験の結果を要約すると以下ようになる．

1. 音弁別の平均反応時間は運動開始から終了に向けて一様にゆっくり減少する．この傾向は，視覚フィードバックの有無によらない．
2. 目標注視のサッカードに先だって，音弁別の反応時間は増加する．
3. 運動終了直前に，反応時間の最小値が階段状に減少する．

新たにわかったことは，弁別課題の反応時間に影響を与える成分には眼球運動に関するものと到達運動に関するものが混じっていること，反応時間の平均値と最小値は異なる挙動を示すことである．

被験者の数が少ないので確定的なことはいえないものの，弁別課題の反応時間が運動のフェーズに応じて様々な変化を見せたという結果は，脳内の計算過程が時間的に交代する複数のプロセスから構成されている可能性を示唆している．すなわち，運動制御の脳内プロセスを考える際は，時間方向に一樣な構造ではなく，運動の進行とともに各プロセスの役割に起伏が生じるような構造を想定する必要があるといえる．

従来の計算理論では，空間方向のモジュール分解は議論されているものの，このような時間方向の役割分担の議論が抜け落ちているように見える．時間方向の処理の流れに関する議論を進展させることにより，軌道計画と運動制御のかかわりを説明する新たな計算モデルが構成できるものと考えている．

4 予測を含めた高次の運動計画

最後に，かつて筆者が提案した高次運動計画の計算モデル [21] について述べる．このモデルは以上の議論に比べてかなり抽象的であるが，上の議論と関連があるのでこの機会に紹介しておきたい．

いま「四角形の板のまわりを指が板からはずれないように注意してなぞる」という課題を考える．この課題を変形された視覚環境で行なうように指示されたとき，被験者はどのような行動をとるであろうか？

心理実験の結果によれば，最初のあいだは小さな運動を繰り返す（速度ピークが複数現れる）が，慣れるに従って一度の運動が大きくなり（速度ピークの数が増える），一気に頂点まで運動するようになる [22]．

このような運動様相の変化は，脳内に，運動指令と結果として現れる到達位置との関係を示す内部モデルが形成された結果生じたと考えられる．それでは，脳はどのようにして内部モデルの変化を運動指令の決

定に使っているのであろうか？

筆者のモデルはこの問題に対する一つの解答を示したものである。アイデアは単純である。まず、運動指令と結果として生じる到達位置との関係を表す確率モデルを用意しておく。この確率モデルは、脳がどの程度の確実さで運動系の性質を同定しているかを表すものである。運動指令選択モジュールは、この確率モデルを用いて手先が頂点を行き過ぎてしまうリスクを計算し、それが一定範囲に収まる運動指令のうち最も運動距離が長いものを選択する。実際に運動を行なったあとは、得られた因果関係を用いて確率モデルを更新する（確率分布が正規分布であると仮定すれば、これは一種のカルマンフィルタになる）。

環境の同定が進んでいないうちは、確率モデルの分散が大きく誤りを犯すリスクを大きめに見込むために、運動距離の小さな指令が選択され、結果として細かい運動を繰り返すことになる。しかし、運動を繰り返すうちに確率モデルの分散が小さくなり、適切な運動指令が選択されるようになるため、一度に大きな運動をするようになる。モデルが示すこのような振舞いは、人間の行動と定性的に一致する。

このモデルの要点は、「未知環境に対するリスクをある程度とった上でパフォーマンス最適化をはかる」という高次の運動計画アルゴリズムを組み立てたことである。このモデルを見て皆さんは何を考えるであろうか？どうすればこのモデルを生理学的知見と対比できるような形に発展させられるだろうか？筆者自身が考える話題をいくつか挙げておきたい。

1. リスクとパフォーマンスのバランスをはかるメカニズムは（仮にあるとすれば）脳の中に埋め込まれているのか？それとも、生まれてからの経験によって獲得されたものか？
2. このモデルでの運動指令選択モジュールは脳のどの部位に相当するのか？上の課題を実行するとき、被験者は意識的に細かい運動を行なうことから、ここでの運動選択は、前節で議論した軌道計画とは異なるレベルの問題である。反応選択という立場にたてば前頭前野かもしれないし、学習プロセスを強化学習と捉える立場からすれば大脳基底核や補足運動野が関与しているかもしれない。この問題は、行動選択と適応の関わりを取り扱っている点で、先に述べた強化学習と教師あり学習の関係と関連している。
3. ここでいう内部モデルは脳のどこにあるのか？小脳外側部なのか、それとも脳の別の部位なのか？
4. 確率モデルを神経モデルにインプリメントするにはどうすればよいか？最も単純な方法は、神経パターンの広がりて表すものであろう。すなわち、活動する神経細胞数が予想の不確定さを表す（つまり、学習が進めば関与する細胞が減りスパースになる）とすればよい。ほかにはどのような方法が考えられるか？
5. このような運動学習モデルと従来の適応学習モデルはどのような関係にあるのか？また、両者を統合することは可能か？そのような統合モデルすることで何が説明できるのか？

確率モデルを用いているという点で、このモデルから分散最小化規範に基づく軌道計画アルゴリズム [23] を思い出す方がいるかもしれない。しかし、分散最小化規範のアイデアは「結果の分散が最小になる運動指令を選択する」ことにある。「結果の分散」をリスクの大きさと読み替えれば、分散最小化はリスク最小化を目指していることになるので、一定リスク下でのパフォーマンス最大化を目指す本モデルとは本質的に異なることがわかる。このほかにも、両者は 1) 対象とする問題（高次の運動選択 vs. 低次の軌道計画）、2) 確率モデルの実体（モデルの不確定性を扱うための仮想的な量 vs. 神経雑音）、3) 運動変化の質（試行錯誤による変化 vs. 神経雑音による変化）の点でも明らかに異なっており、別々の問題を扱うモデルであると考えべきだろう。

5 むすび

小脳、予測、適応といったキーワードを手がかりに、到達運動に関わるいくつかの問題について議論してきた。以上の議論に対する質問、コメント、反論が多数いただけることを期待している。

なお、本研究の一部は、科学技術振興事業団・戦略基礎研究事業「脳を創る」の「感覚運動統合理論に基づく『手と脳』の工学的実現」（代表者：石川正俊）および「運動の学習制御における小脳機能の解明」（代表者：永雄総一）により行なっているものである。

参考文献

- [1] 虫明 元: 基底核, 小脳と大脳皮質の機能連関. 神経進歩, 39, 277–289, 1995.
- [2] Kawato M, Gomi H: The cerebellum and VOR/OKR learning models. Trends Neurosci, 15, 445–453, 1992.
- [3] Schweighofer N, Arbib MA, Kawato M: Role of the cerebellum in reaching movements in humans. I. Distributed inverse dynamics control. Eur J Neurosci, 10, 86–94, 1998.
- [4] Schweighofer N, Spoolstra J, Arbib MA, Kawato M: Role of the cerebellum in reaching movements in humans. II. A neural model of the intermediate cerebellum. Eur J Neurosci, 10, 95–105, 1998.
- [5] Krakauer JW, Ghilardi MF, Ghez C: Independent learning of internal models for kinematic and dynamic control of reaching. Nature Neurosci, 2, 1026–31, 1999.
- [6] Graziano MS, Yap GS, Gross CG: Coding of visual space by premotor neurons. Science, 266, 1054–57, 1994.
- [7] Doya K: What are the computations of the cerebellum, the basal ganglia and the cerebral cortex. Neural Networks, 1999.
- [8] Thach WT, Mink JW, Goodkin HP, Keating JG: Combining versus gating motor programs: Differential roles for cerebellum and basal ganglia? in Roles of the cerebellum and basal ganglia in voluntary movement, Eds. Mano N, Hamada I, DeLong MR, Elsevier, 235–245, 1993.
- [9] Jacobs RA, Jordan MI, Nowlan SJ, Hinton GE: Adaptive mixtures of local experts. Neural Computation, 3, 79–87, 1991.
- [10] Kawato M, Wolpert D: Internal models for motor control. Novartis Foundation Symposium, 218, 291–307, 1998.
- [11] Kurata K, Hoshi E: Reacquisition deficits in prism adaptation after muscimol microinjection into the ventral premotor cortex of monkeys. J Neurophysiol, 81, 1927–38, 1999.
- [12] 赤司裕一, 阪口 豊: 視覚運動変換におけるランプ型適応とステップ型適応の違い. 電子情報通信学会技術研究報告, NC99-174, 2000.
- [13] Bizzi E, Accornero N, Chapple W, Hogan N: Posture control and trajectory formation during arm movement. J Neurosci, 4, 2738–44, 1984.
- [14] Houk JC, Wise SP: Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. Cerebral Cortex, 5, 95–110, 1995.
- [15] Prablanc C, Pelisson D, Goodale MA: Visual control of reaching movements without vision of the limb. I. Role of retinal feedback of target position in guiding the hand. Exp Brain Res, 62, 293–302, 1986.
- [16] Pelisson D, Prablanc C, Goodale MA, Jeannerod M: Visual control of reaching movements without vision of the limb. II. Evidence of fast unconscious processes correcting the trajectory of the hand to the final position of a double-step stimulus. Exp Brain Res, 62, 303–311, 1986.
- [17] Bard C, Turrell Y, Fleury M, Teasdale N, Lamarre Y, Martin O: Deafferentation and pointing with visual double-step perturbations. Exp Brain Res, 125, 410–416, 1999.
- [18] Pashler H: Dual-task interference in simple tasks: Data and theory. Psychological Bulletin, 116, 220–244, 1994.
- [19] Els JR: Analysis of temporal and attentional aspects of movement control. J Exp Psychol, 99, 10–21, 1973.
- [20] Ketelaars MAC, Khan MA, Franks IM: Dual-Task Interference as an Indicator of On-Line Programming in Simple Movement Sequences. J Exp Psychol:HPP, 25, 1302–1315, 1999.
- [21] 阪口 豊: 内部モデルの信頼度に基づく運動計画のアルゴリズム. 電子情報通信学会論文誌, J79-D-II, 248–256, 1996.
- [22] Akamatsu M: The influences of combined visual and tactile information on finger and eye movements during shape tracing. Ergonomics, 35, 647–660, 1992.
- [23] Harris CM, Wolpert DM: Signal-dependent noise determines motor planning, Nature, 394, 780–784, 1998.

3日目 報酬の予測 大脳基底核

コーディネータ：小池康晴，鮫島和行

1. ねらい

学習の枠組みの中でも報酬というスカラーだけから学習する理論的枠組みである強化学習は，大脳基底核の機能との整合性から，これまで脳の実験家と理論家の相互の交流によってさまざまな実験や理論の展開が進んでいる分野の1つである．3日目の報酬予測と大脳基底核では，この分野の理論的基礎を学び，また，現在進んでいる理論と実験の協力による脳の実験の実際がどのように進んでいるのかを概観する．

2. 講義内容・関連

午前 強化学習の基礎と演習（小池・鮫島）

報酬に基づいて行動を生成する理論として強化学習がある．強化学習の基礎についての講義および演習を行う．特に，報酬の予測を行う価値関数の概念や，報酬を最大化する最適行動の生成や最適価値関数の学習はどのようなアルゴリズムによって実現されるか？報酬の予測や行動の生成に状態の予測はいかにして関係するか？などの講義を行う．

午後 大脳基底核の報酬関連活動（彦坂・中原）

実際に動物が報酬にもとづいた行動を生成する際の神経細胞活動，特に大脳基底核の神経活動についての講義を行う．特に報酬および報酬予測に対して反応するドーパミン細胞の発火が大脳基底核の入力部である線条体細胞にいかにか影響するのかや，さらに，報酬予測の理論的枠組みである強化学習と，実験的事実である大脳基底核および運動関連大脳皮質の活動をいかに結び付け，環境の変化，報酬の変化，行動の生成に関して脳の情報処理をいかにして理解するかの方法についての講義する．

3. 必読文献

R.S.Sutton and A.G.Barto, Reinforcement learning: An introduction, Chapter 3. The Reinforcement learning problem, pp.51-85, 1998

A.G. Barto, Adaptive Critics and the Basal Ganglia ,In Models of Information Processing in the Basal Ganglia, James C. Houk , Joel L. Davis , David G. Biser , pp. 215-232 1995

Nakahara, H., Doya, K., Hikosaka, O. (1998) Benefit of multiple representations for motor sequence control in the basal ganglia loops. BSIS Tech. Rep. No.98-5

強化学習の基礎

小池 康晴[†], 鮫島 和行[‡]

[†] 東京工業大学 精密工学研究所

[‡] 科学技術振興事業団 ERATO 川人学習動態脳プロジェクト

1 教師無し学習

赤ちゃんが生れてから、手を動かしたり、物をつかんだりするとき、手の動かし方や、物のつかみ方を直接教えてもらうことなく、環境との相互作用により、自然に学習する。このような学習は、教師なし学習とよばれている。一方、絵や文字を覚えるような学習は、覚える絵や文字が与えられているため、教師あり学習と呼ばれる。しかし、これらは、完全に分離することはできない。例えば、言葉を覚えるとき、自分の発生した音と他人が発生した音は、直接比較することができるため、その差をつかって教師あり学習をすることができる。しかし、調音器官の筋肉一本一本の動きは直接知ることができないため、教師無し学習を行うことになる。

近年、目標パターン（この例では言葉）は与えられずに、出力が良かったか、悪かったかだけ（この例では、上手く発音できたらほめられる）を与えるだけで、その評価を最大にする出力を学習する枠組みとして強化学習（Reinforcement Learning）が提案されている。

2 強化学習の要素

まずはじめに、強化学習の枠組みを簡単に説明する。強化学習では、様々な行動を試してみて、より良い報酬が得られる行動を選択することを行う。

- policy

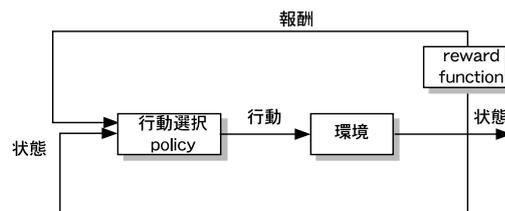
行動を決めるルール

- reward function

ある状態において獲得できる報酬

- value function

将来にわたって獲得できる報酬の総和



強化学習の他の学習則と異なる特徴は、正しい行動を教えられるのではなく、取った行動を評価することで学習をおこなうことである。そのために、二つのプロセスを繰り返しおこなう。一つは、現在のポリシーに従って行動をおこない、得られた報酬から価値関数 (Value function) を作成すること (ポリシーの評価) であり、もう一つは、現在の価値関数から最適なポリシーを作成すること (ポリシーの改善) である。

3 環境との相互作用による学習

強化学習では、環境から得られる最終的な累積報酬を最大化することで学習を行う。累積報酬は、以下の式で与えられる。

$$R_t = \sum_{k=0}^T \gamma^k r_{t+k+1}$$

ここで、 T は、最終時刻、 γ は、遠い将来に得られる報酬ほど割引いて評価するための割引き率 (discount factor) であり、 $0 \leq \gamma \leq 1$ である。

3.1 マルコフ性

ここで、マルコフ性の定義をおこなう。数学的記述の簡単化のために状態と、報酬の数は有限であると仮定する。時刻 t で取った行動に対して、時刻 $t+1$ において、どのように応答するかを考える。一般的には、時刻 $t+1$ 以前に起こった全ての事象が関係するため、次のように定義される。

$$Pr\{s_{t+1} = s', r_{t+1} = r | s_t, a_t, r_t, s_{t-1}, a_{t-1}, \dots, r_1, s_0, a_0\}$$

一方、状態信号がマルコフ性を持つならば、 $t+1$ の応答は一時刻前 t の状態と行動だけによって決る。

$$Pr\{s_{t+1} = s', r_{t+1} = r | s_t, a_t\}$$

このように、マルコフ性を仮定すると、現在の状態と行動から次の時刻の状態と報酬を予測することができる。さらに、繰り返し計算により、すべての将来の状態と報酬を予測することができる。

マルコフ性を満足する強化学習は、マルコフ決定過程 (Markov decision process:MDP) と呼ばれる。有限 MDP では、任意の状態 s と行動 a が与えられると、可能な次の状態 s' の確率は、次の式で与えられる。

$$\mathcal{P}_{ss'}^a = Pr\{s_{t+1} = s' | s_t = s, a_t = a\}$$

これは、遷移確率 (transition probabilities) と呼ばれる。また、現在の状態 s と行動 a が与えられたとき次の状態 s' での報酬の期待値は、以下のようになる。

$$\mathcal{R}_{ss'}^a = E\{r_{t+1} | s_t = s, a_t = a, s_{t+1} = s'\}$$

3.2 価値関数 (value function)

強化学習では、報酬を評価してその評価を最大化することで学習を行う。ここでは、現在の状態 (あるいは、行動) がどのくらい良いのか、を計る関数として、価値関数というものを考える。“どのくらい良いのか” というこを、将来にわたって得られる報酬によって定義する。

方策 π というのは、状態 $s \in \mathcal{S}$ で行動 $a \in \mathcal{A}(s)$ をとることであり、 $\pi(s, a)$ と表す。方策 π のもとで、状態 s の価値は、以下のように定式化できる。

状態価値関数 (state-value function for policy π)

$$V^\pi(s) = E_\pi\{R_t | s_t = s\} = E_\pi\left\{\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s\right\}$$

同様に、方策 π のもとで、状態 s において行動 a を取ることの価値は、以下のように定義できる。

行動価値関数 (*action-value function for policy π*)

$$Q^\pi(s, a) = E_\pi\{R_t | s_t = s, a_t = a\} = E_\pi\left\{\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s, a_t = a\right\}$$

価値関数は、以下のような再帰的な関係をもっている。 *Bellman equation*

$$\begin{aligned} V^\pi(s) &= E_\pi\{R_t | s_t = s\} \\ &= E_\pi\left\{\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s\right\} \\ &= E_\pi\left\{r_{t+1} + \sum_{k=0}^{\infty} \gamma^k r_{t+k+2} | s_t = s\right\} \\ &= \sum_a \pi(s, a) \sum_{s'} \mathcal{P}_{ss'}^a \{R_{ss'}^a + \gamma E_\pi\left\{\sum_{k=0}^{\infty} \gamma^k r_{t+k+2} | s_{t+1} = s'\right\}\} \\ &= \sum_a \pi(s, a) \sum_{s'} \mathcal{P}_{ss'}^a \{R_{ss'}^a + \gamma V^\pi(s')\} \end{aligned}$$

この式は、すべての行動に対して、すべての期待される次の状態での（ディスカウントされた）価値と報酬の和を生起確率で重みづけしたものである。

3.3 Optimal Value Functions

評価することが出来るようになったため、次に行うことは、最適な価値関数を求めることである。これは、もっとも報酬を多く得られる方策を求めることである。

最適な状態価値関数を V^* で表すと、

$$V^*(s) = \max_{\pi} V^\pi(s)$$

となる。また、最適な行動価値関数を Q^* で表すと、

$$Q^*(s, a) = \max_{\pi} Q^\pi(s, a)$$

となる。これは、状態と行動のペア (s, a) に対して、状態 s において、行動 a をとり、以後、最適な方策に従って行動したときに得られる報酬の期待値である。したがって、 Q^* は、 V^* を用いて以下の様を書くことができる。

$$Q^*(s, a) = E_\pi\{r_{t+1} + \gamma V^*(s_{t+1}) | s_t = s, a_t = a\}$$

V^* に対する最適な Bellman 方程式は、以下のようにかける。

$$\begin{aligned} V^*(s) &= \max_{a \in A(s)} Q^{\pi^*}(s, a) \\ &= \max_a E_{\pi^*}\{R_t | s_t = s, a_t = a\} \\ &= \max_a E_{\pi^*}\left\{\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s, a_t = a\right\} \\ &= \max_a E_{\pi^*}\left\{r_{t+1} + \sum_{k=0}^{\infty} \gamma^k r_{t+k+2} | s_t = s, a_t = a\right\} \\ &= \max_a E\{r_{t+1} + \gamma V^*(s_{t+1}) | s_t = s, a_t = a\} \\ &= \max_a \sum_{s'} \mathcal{P}_{ss'}^a \{R_{ss'}^a + \gamma V^*(s')\} \end{aligned}$$

また、 Q^* に対する最適な Bellman 方程式は、以下のようである。

$$\begin{aligned} Q^*(s, a) &= E\{r_{t+1} + \gamma \max_{a'} Q^*(s_{t+1}, a') | s_t = s, a_t = a\} \\ &= \sum_{s'} P_{ss'}^a \left(R_{ss'}^a + \gamma \max_{a'} Q^*(s', a') \right) \end{aligned}$$

4 Temporal-Difference Learning

TD 学習は、Monte Carlo 法と Dynamic Programming の組合せと考えることができる。環境のモデルを使わず経験的に学習をおこなう、という点で MC 法に、また、最終結果を待たずに、評価を途中で更新する点で DP 法に似ている。

Monte Carlo 法では、各時刻の報酬が分るまで待ち価値関数を更新するため、単純な Monte Carlo 法では、価値関数の更新は次のようになる。

$$V(s_t) \leftarrow V(s_t) + \alpha [R_t - V(s_t)]$$

一方 TD 法では、次のステップを待つだけで、価値関数を更新する。

$$V(s_t) \leftarrow V(s_t) + \alpha [r_{t+1} + \gamma V(s_{t+1}) - V(s_t)]$$

$r_{t+1} + \gamma V(s_{t+1}) - V(s_t)$ は、TD 誤差と呼ばれている。

4.1 Q-learning

off-policy TD \rightarrow Q-learning

各状態において、可能な行動の中で最も行動評価関数の値が高い行動をとるように学習を行う方法を Q Learning と呼ぶ。その学習は、

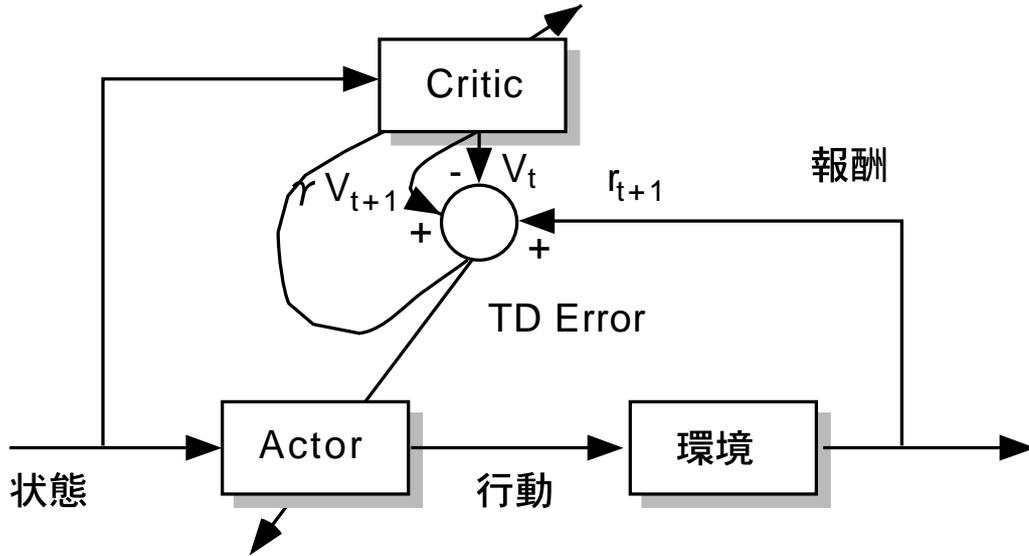
$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_a Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$$

というように行われる。Q Learning が方策 off 型と呼ばれるのは、方策に関係なく行動価値関数の最大値で行動価値関数を更新するためである。

4.2 Actor-Critic

on-policy TD \rightarrow Actor-Critic

Actor-Critic 法は、価値関数とは独立に、方策を表現する構造を別に持っている。これは、行動を選択するために用いられるため “actor” と呼ばれている。また、価値関数を予測する部分は、actor によって選ばれた行動を批判するために “critic” と呼ばれている。



状態価値関数は、

$$V(s_t) \leftarrow V(s_t) + \alpha[r_{t+1} + \gamma V(s_{t+1}) - V(s_t)]$$

によって更新される。

TD 誤差は、行った行動を評価する為に用いられる。TD 誤差が正の値ならば、行った行動は価値を高めたことになるので、より選択されるように、逆に、TD 誤差が負の値ならば、選択されないようにするほうがよい。

$p(s, a)$ を actor が状態 s で行動 a を取る確率だとすると、行動選択は、

$$\pi_t(s, a) = Pr\{a_t = a | s_t = s\} = \frac{e^{p(s,a)}}{\sum_b e^{p(s,b)}}$$

で表される。

先に示したように actor は、

$$p(s_t, a_t) \leftarrow p(s_t, a_t) + \beta \delta_t$$

として、学習をおこなう。ここで、 β は正のステップサイズ変数であり、 δ_t は TD 誤差である。

4.3 TD(λ)

これまでみてきた TD 法は、1 ステップ後での評価を比較して価値関数を更新していた。ここでは、 n ステップ後まで考えるように拡張をおこなう。

n ステップ後までの報酬を考慮にいと、最終的に得られる累積報酬は以下ようになる。

$$R_t^{(n)} = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots + \gamma^{n-1} r_{t+n} + \gamma^n V_t(S_{t+n})$$

さらに、任意の n ステップの報酬だけでなく、 n ステップの報酬の平均値を考えることもできる。

TD(λ) は、 n ステップのバックアップを平均化する方法の一つである。

$$R_t^\lambda = (1 - \lambda) \sum_{n=1}^{\infty} \lambda^{n-1} R_t^{(n)}$$

ここでは、 $\lambda^{n-1}, 0 \leq \lambda \leq 1$ によって重み付けされている。 $(1 - \lambda)$ は、正規化のための項である。

実際にアルゴリズムを実装するためには、逆の見方をするほうが分かりやすい。そこで、*eligibility trace* という考え方を導入する。ある時刻 t で状態 s のとき、*eligibility trace* $e_t(s)$ は、以下のように定義できる。

$$e_t(s) = \begin{cases} \gamma \lambda e_{t-1}(s) & \text{if } s \neq s_t; \\ \gamma \lambda e_{t-1}(s) + 1 & \text{if } s = s_t; \end{cases}$$

これを用いると、価値関数の更新式は以下ようになる。

$$\Delta V_t(s) = \alpha \delta_t e_t(s), \text{ for all } s \in \mathcal{S}$$

5 状態予測モデルと価値関数の学習

ここまでは、環境がどのように変化するかを予測する状態予測モデルを用いずに経験から直接価値関数を学習し、その価値関数にしたがって行動を決定するか、または行動を決定する機構を価値関数を学習するための誤差信号を用いて学習する (Actor critic) を紹介した。ここでは、このようなモデルを持たない (model free) 強化学習に対して、モデルを持った場合 (model based) の強化学習についていくつかの手法を紹介する。

5.1 状態価値関数を用いた off-policy method

Q-learning では常に最大の行動価値を選択したという仮定で価値関数を更新するが、Actor-critic では実際の経験に基づく状態価値の時間差のみで更新する。この意味で Actor-Critic は on-policy であり、Q-learning は off-policy の更新法であるといえる。Actor critic のように状態価値関数を用いた強化学習では、次の状態を予測するモデルを用いることによって Q-learning と同様に off-policy の更新を行うことができる。

すなわち、状態 $s(t)$ に関する状態評価値 $V(s)$ の更新則に用いる TD 誤差を行動 a を選択した時の 1step 先の状態の予測 $s(t+1) = s_{\text{model}}(s(t), a)$ と報酬の予測 $\hat{r}(s(t), a) = r_{\text{model}}(s(t), a)$ を用いることによって擬似的な TD 誤差

$$\hat{\delta}(s(t), a) = \hat{r}(t) + \gamma V(\hat{s}(t+1)) - V(s(t))$$

を求める。この $\hat{\delta}(s, a)$ の a に関する最大値を用いて価値関数を更新する。

$$V(s) \leftarrow V(s) + \eta \max_{a \in A} \hat{\delta}(s, a)$$

ここで η は学習係数。

また、行動選択に際してもモデルと価値関数から行動に導くことができる。行動価値関数 $Q(s, a)$ を用いた場合、現在近似している価値をより大きくする方向に行動する greedy policy

$$a' = \arg \max_{a \in A} Q(s, a)$$

とするか、または行動価値にしたがって確率をボルツマン分布にしたがって与える

$$P(a = a' | s) = \frac{\exp(\beta Q(s, a'))}{\sum_{a \in A} \exp(\beta Q(s, a))}$$

とすることによって価値関数を学習するのみでポリシーが得られる。ここで β は行動選択のランダムネスを決定するパラメータである。

これに対して状態評価関数のみを持ち、Actor-Critic のように行動を決定する別の構造をもたない場合には、greedy policy は1ステップ先の状態価値と報酬が最大になるように行動を選択するので。

$$a' = \arg \max_{a \in A} (\hat{r}(s(t)) + \gamma V(\hat{s}(t+1))) \quad (1)$$

$$= \arg \max_{a \in A} \hat{\delta}(s(t), a) \quad (2)$$

同様に価値に従った確率で行動する場合

$$P(a = a' | s) = \frac{\exp(\beta \hat{\delta}(s, a'))}{\sum_{a \in A} \exp(\beta \hat{\delta}(s, a))}$$

のように予測モデルによって1ステップ先の価値との差によって行動を決定することができる。

5.2 モデルを用いた価値関数学習の高速化: Dyna

これまで説明してきた強化学習では、経験(状態と行動と評価値の対)から直接行動(policy)を更新する、または価値関数(value)を更新するという手法を用いている。

経験をまずモデルの更新に用い、そのモデルから得られる擬似的な経験を用いて価値関数を学習をする間接的な学習法を考える。Dynaは、経験からモデルの学習と行動価値関数の更新の両方を行い、さらにモデルからの擬似的経験によって価値関数の更新を高速化するというアーキテクチャである。

Dynaでは、モデルから得られる擬似的経験による更新はモデルの持つ経験の中からランダムに選択される。価値関数の学習では、ランダムに擬似経験を与えるよりも、より大きなTD誤差が発生した状態にたいして遷移する状態を優先して学習させるほうがより強力に価値関数の学習を高速化できるはずである。Prioritized sweepingという考え方は、このような重要な状態に優先順位をつけ、その順に擬似経験を生成して価値関数を更新する手法である。実際の経験に基づくTD誤差が大きな状態に高い優先順位をつけ、さらに、その状態へ遷移する状態をモデルから推定して優先順位を順に時間的に逆方向にわりつけるという手法である。

5.3 最近の研究

最近複雑な環境を簡単な環境に分割し、それぞれの部分問題を解いた後に統合するモジュール強化学習や階層構造を導入することによってより現実的な速さで学習を行おうとする試みがされている。

Wolpert and Kawato(1998)によって提案されたmultiple pared forward inverse model(MPFIM)は、複雑な環境を簡単な複数の予測モデルによって部分的に近似させ、それぞれの予測モデルの予測の正確さによって複雑な環境を分割する手法を提案している。この複数の予測モデルを用いたモデルベース強化学習(multiple model based reinforcement learning, MMRL)を紹介する。

6 脳のモデル

近年、サルの場合条件反射学習の実験において、大脳基底核のドーパミンニューロンは、学習前は、報酬そのものに反応するが、学習が進むにつれて、報酬を予測させる刺激に対して反応するようになることが示された。

この反応は、TD誤差の振舞いと良く似ているため、強化学習で用いられる、状態価値関数や、行動価値関数のような出力を持つ細胞が脳のなかに獲得されている可能性もある。

Translation of motivation into action in the basal ganglia

Okihide Hikosaka, Reiko Kawagoe, and Yoriko Takikawa

e-mail: hikosaka@med.juntendo.ac.jp

Dept. of Physiology, Juntendo University, School of Medicine, 2-1-1 Hongo, Bunkyo-ku, Tokyo
113-0033, Japan

ABSTRACT

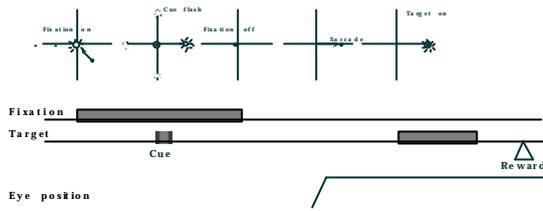
The basal ganglia, especially the ventral striatum, have been implicated in control of action based on motivation^{1,2,3,4}. A prevalent view is that nigro-striatal dopaminergic neurons carry reinforcement signals to modulate the cortico-striatal signal transmissions^{5,6,7}. However, it is still unknown how such reinforcement signals affect the output of the striatum in relation to behavior. To answer this question, we devised a memory-guided saccade task in which only one out of four directions was rewarded, and examined single cell activity in the caudate nucleus. We found that visual or memory-related responses of presumed projection neurons in the caudate were frequently modulated by expectation of reward, either as an enhancement or as a reduction of response. The cell's preferred direction often changed with the change in the rewarded direction, implying a short-term synaptic plasticity. The modulation of caudate cell activity was correlated with changes in saccade parameters. Our results suggest that the caudate contributes to the determination of oculomotor outputs by affiliating motivational values to visual information.

INTRODUCTION

We used two male Japanese monkeys (*Macaca fuscata*). Under general anesthesia, we implanted a head holder, chambers for unit recording, and a scleral search coil²¹. The monkeys were trained to perform saccade tasks, especially a memory-guided saccade task²⁷. Eye movements were recorded using the search coil method. We recorded extracellular spike activity of presumed projection neurons which showed very low spontaneous activity²⁸, but not of presumed interneurons which showed irregular tonic discharge²⁹. For each cell that showed visual or memory-related responses, we used a set of four target locations with the same eccentricity that were arranged in either normal or oblique angles, depending on the cell's receptive field. The recording sites were verified using MRI (Hitachi, AIRIS, 0.3T).

The monkeys performed the memory-guided saccade task in two different reward conditions: all-directions-rewarded condition (ADR) and one-direction-rewarded condition (1DR). For every caudate cell recorded, we required the monkeys to perform one block of ADR and four blocks of 1DR (i.e., four different rewarded directions).

METHODS



In both conditions, a task trial started with onset of a central fixation point. While the monkeys were fixating the fixation point, a cue stimulus whose location must be remembered was presented randomly at one of the four directions. After 1-1.5 s, the fixation point turned off, and the monkeys were required to make a saccade to the previously cued location.

In ADR, every correct saccade was rewarded with a liquid reward together with a tone stimulus. In IDR, an asymmetric reward schedule was used in that only one of the four directions was rewarded while the other directions were either not rewarded (exclusive IDR) or rewarded with a smaller amount (about 1/5) (relative IDR). The highly rewarded direction was fixed in a block of experiments which included 60 successful trials. Even for the non-rewarded or less-rewarded direction, the monkeys had to make a correct saccade. The correct saccade was indicated by a tone stimulus with no or small reward, which was followed by the next trial; if the saccade was incorrect, the same trial was repeated. The amount of reward per trial was set approximately the same between IDR and ADR. The target cue was chosen pseudo-randomly such that the four directions were randomized in every sub-block of four trials; thus, one block of experiment (60 trials) contained 15 trials for each direction. IDR was performed in four blocks, in each of which a different direction was rewarded highly. Other than the actual reward, no indication was given to the monkeys as to which direction was currently rewarded.

For each cell responding to the cue stimulus, we first determined the duration of the response (test duration) based on cumulative time histograms, usually based on the most robust response. A control duration (usually 500 ms) was set just before the onset of the fixation point. The cell's response was calculated, for each trial, as the spike frequency during the test duration minus the spike frequency during the control duration.

RESULTS AND DISCUSSION

We trained two monkeys to perform a memory-guided saccade task in two reward conditions: all-directions-rewarded condition (ADR) and one-direction-rewarded condition (IDR). In ADR, which is a conventional reward schedule, the monkeys were rewarded each time they made a memory-guided saccade correctly. In IDR, which we devised specifically for the present study, the monkeys were rewarded when the cue stimulus was presented in one particular direction out of four and the saccade was made correctly; they were not rewarded (exclusive IDR) or rewarded with a smaller amount (relative IDR) for the other three directions, but had to make a correct saccade to proceed to the next trial. The rewarded direction was fixed in a block of 60 trials, and a total of four blocks was performed with four different rewarded directions. Thus, the cue stimulus had two meanings: (1) the direction of the saccade to be made later, and (2) whether or not a big reward was to be obtained after the saccade.

Among 241 cells we recorded in the caudate nucleus, there were cells showing phasic visual

responses to the cue stimulus (n=114), sustained activity during the delay period (memory-related response) (n=79), saccadic responses (n=92), and activity preceding the cue stimulus (n=89). In this report, we concentrate on 87 cells with visual or memory-related responses in which 4 blocks of 1DR and 1 block of ADR were fully examined. We defined a visual response to be phasic activity that started within 200 ms after onset of the cue stimulus and a memory-related response to be sustained activity that started 200 ms after the cue onset and ended before or with the saccade. Among them, 27 out of 45 cells (60 %) with visual response and 20 out of 50 cells (40 %) with memory-related response showed clear direction selectivity when tested in ADR (one-way ANOVA (cued direction), $P < 0.01$) (c.f., note that the two types of response could be observed in a single cell). The preferred direction was usually contralateral (70 %), as reported previously⁸.

We found, however, that such spatial selectivity depended on the reward condition. A typical cell is shown in Fig. 1, which was recorded in the right caudate nucleus. In ADR, it responded to the left (contralateral) cue stimulus most vigorously, while the response to the right cue was meager. The cell's direction selectivity is shown at top as a polar diagram.

In 1DR, however, the cell's direction selectivity changed completely. For example, when the rewarded direction was right, the cell responded to the right cue stimulus much better than to the other directions. Accordingly, the cell changed its preferred direction in different blocks such that the response was greatest for the rewarded direction. The response was clearly dependent on the reward condition [two-way ANOVA (reward condition x cued direction), main effect of reward condition: $F(1, 181) = 689.243$; $P < 0.0001$].

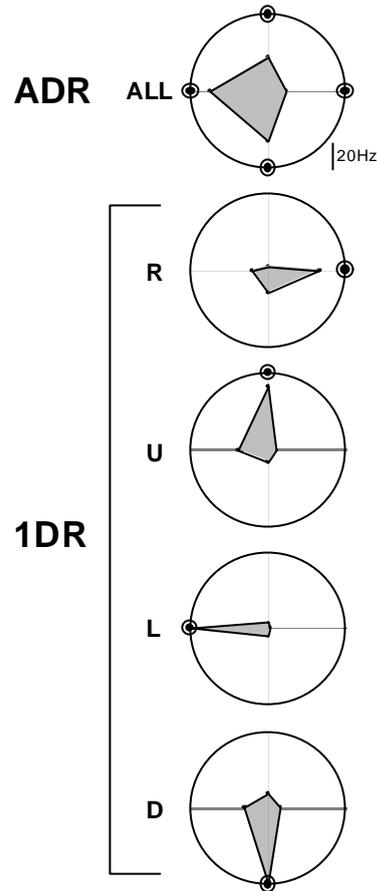


Figure 2:

The caudate cell shown in Fig. 2 was also dependent on reward expectation, but in the opposite manner. In ADR, the cell showed virtually no response to any of the four cue stimuli. In 1DR, however, it showed vigorous responses to the cue that indicated no reward, while it showed no response to the rewarded cues, no matter which direction was rewarded.

The cells shown in Fig. 1 and 2 were not exceptional ones. As shown in Fig. 3A, most caudate cells showed either a strong enhancement (data points close to the ordinate) or a reduction (data points close to the abscissa) of response by expectation of reward. A statistically significant modulation was found in 76 out of 87 cells (87 %) in either the visual or memory-related response: visual response, 36/45 (80 %); memory response, 43/50 (86 %) [two-way ANOVA (reward condition x cued direction), main effect of reward condition; $P < 0.01$]. Among the 76 modulated

cells, 64 cells (visual: 31, memory: 36) showed an enhancement ('reward-facilitated cells'), while 12 cells (visual: 5, memory: 7) showed a reduction of response ('reward-suppressed cells'). Similar results were obtained using the exclusive IDR and relative IDR.

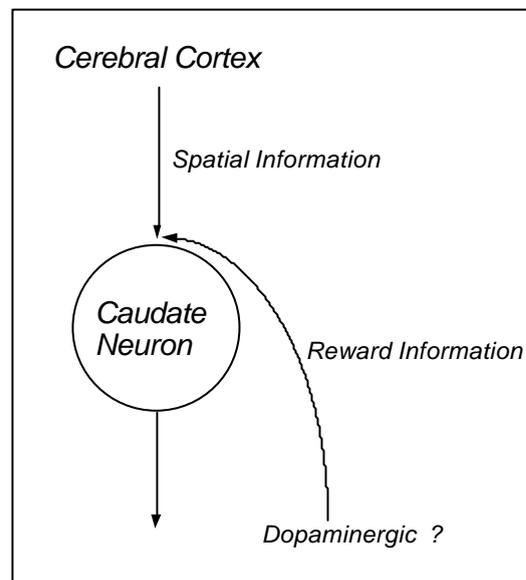
That the monkeys were more motivated when reward was expected was indicated in the changes in saccade parameters. The latencies were shorter (Fig. 3B) and the peak velocities were higher (Fig. 3C) when the saccades were followed by reward than when they were not (paired t-test, $P < 0.0001$).

We then asked how the caudate cells changed their response when the rewarded direction was changed (Fig. 4). In the first block of IDR for the reward-facilitated cell (shown in Fig. 1), the rewarded direction was left, which was the cell's preferred direction in ADR (Fig. 4A, left). The responses were initially strong for all directions except for right, but the responses to the left cue gradually increased, while the responses to the other cues decreased rapidly and stayed close to zero. In the next block (Fig. 4A, right), the rewarded direction was changed to right which was the non-preferred direction in ADR. Again, the responses were initially strong for all directions, but decreased gradually while only the response to the right cue survived. The time course for the reward-suppressed cell (shown in Fig. 2) was quite opposite to that of the reward-facilitated cell shown in Fig. 4A. For each block, the cell initially showed almost no response to any direction, but then started responding to the three directions that indicated no reward (Fig. 4B).

A similar time course of response modulation was observed in the other reward-contingent caudate cells, especially for the non-rewarded cues. Specifically, among 64 reward-facilitated cells, 27 decreased their response while the others showed no significant change; among 12 reward-suppressed cells,

4 increased their responses while the others showed no change [unpaired t-test (comparison between the initial 15 trials and the following trials, $P < 0.01$).

Neurons that we recorded had low spontaneous activity and were presumably projection neurons which are GABAergic⁹. They are thought to modulate the final inhibitory outputs of the basal ganglia, either by disinhibition or by enhancement of inhibition^{10,11,12}. Anatomically, the striatal projection neurons are characterized by numerous spines on their dendrites^{13,14} to which glutamatergic cortico-striatal axons and dopaminergic axons make synaptic contacts^{15,16}. Schultz and his colleagues have demonstrated that dopaminergic neurons in the substantia nigra show responses to sensory stimuli that predict the upcoming reward^{17,7}. Thus, a caudate neuron could receive spatial information via the cortico-striatal inputs¹⁸ and reward-related information via the dopaminergic input¹⁷.



Based on these considerations, we propose that the efficacy of the cortico-striatal synapses would be enhanced or depressed depending on the combination of these two inputs. In reward-facilitated cells (as shown in Fig. 1), the co-activation of these two inputs

would lead to synaptic enhancement, while activation of either one of them alone would lead to depression. The scenario would be opposite in the case of reward-suppressed cells (as shown in Fig. 2). Different dopaminergic receptors, such as D1 and D2, might be involved in such excitatory and inhibitory processes¹⁹. These mechanisms, in fact, have been suggested in relation to long-term depression and long-term potentiation²⁰. The synaptic plasticity in our case would be a short-term one, because the preferred direction changed fairly rapidly in a block of IDR trials.

The reward-contingent modulation of caudate cell activity was correlated with the changes in saccade latency and velocity. A mechanism underlying the changes may be the serial inhibitory connections from the caudate to the superior colliculus through the substantia nigra pars reticulata^{11,21}. An enhancement of caudate cell activity when reward is expected (as in Fig. 1) would lead to an enhanced disinhibition of the superior colliculus and consequently a reduction of saccade latency and an increase in saccade velocity, especially for memory-guided saccades²², which we observed in the present study. On the other hand, an enhancement of caudate cell activity when reward was not expected (as in Fig. 2) might affect the so-called indirect pathway (including the globus pallidus external segment²³ and subthalamic nucleus²⁴), which would lead to the suppression of saccades to the non-rewarded cues, as observed in our study. The above scheme, however, needs to be examined in future studies.

It has been suggested that the basal ganglia contribute to the selection of action^{25,26}. Our study suggests that a critical determinant for the selection is expectation of reward (or motivation). The caudate nucleus, part of the dorsal striatum, would play an important role in such a decision-making process.

References

1. Mogenson, G.J., Jones, D.L. & Yim, C.Y. From motivation to action: functional interface between the limbic system and the motor system. *Progress in Neurobiology* **14**, 69-97 (1980).
2. Robbins, T.W. & Everitt, B.J. Neurobehavioural mechanisms of reward and motivation. *Curr. Opin. Neurobiol.* **6**, 228-236 (1996).
3. Schultz, W., Apicella, P., Scarnati, E. & Ljungberg, T. Neuronal activity in monkey ventral striatum related to the expectation of reward. *J. Neurosci.* **12**, 4595-4610 (1992).
4. Bowman, E.M., Aigner, T.G. & Richmond, B.J. Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *J. Neurophysiol.* **75**, 1061-1073 (1996).
5. Houk, J.C., Adams, J.L. & Barto, A. in *Models of information processing in the basal ganglia* (eds. Houk, J.C., Davis, J.L. & Beiser, D.G.) 249-270 (MIT Press, Cambridge, MA, 1995).
6. Wickens, J. & Kotter, R. in *Models of information processing in the basal ganglia* (eds. Houk, J.C., Davis, J.L. & Beiser, D.G.) 187-214 (MIT Press, Cambridge, MA, 1995).
7. Schultz, W., Dayan, P. & Montague, P.R. A neural substrate of prediction and reward. *Science* **275**, 1593-1599 (1997).
8. Hikosaka, O., Sakamoto, M. & Usui, S. Functional properties of monkey caudate neurons. II. Visual and auditory responses. *J. Neurophysiol.* **61**, 799-813 (1989).
9. Ribak, C.E., Vaughn, J.E. & Roberts, E. The GABA neurons and their axon terminals in rat corpus striatum as demonstrated by GAD

- immunocytochemistry. *J. Comp. Neurol.* **187**, 261-284 (1979).
10. Chevalier, G. & Deniau, J.M. Disinhibition as a basic process in the expression of striatal functions. *Trends Neurosci.* **13**, 277-280 (1990).
11. Hikosaka, O. & Wurtz, R.H. in *The Neurobiology of Saccadic Eye Movements* (eds. Wurtz, R.H. & Goldberg, M.E.) 257-281 (Elsevier, Amsterdam, 1989).
12. Alexander, G.E. & Crutcher, M.D. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci.* **13**, 266-271 (1990).
13. Preston, R.J., Bishop, G.A. & Kitai, S.T. Medium spiny neuron projection from the rat striatum: an intracellular horseradish peroxidase study. *Brain Res.* **183**, 253-263 (1980).
14. Kawaguchi, Y., Wilson, C.J. & Emson, P.C. Projection subtypes of rat neostriatal matrix cells revealed by intracellular injection of biocytin. *J. Neurosci.* **10**, 3421-3438 (1990).
15. Smith, A.D. & Bolam, J.p. The neural network of the basal ganglia as revealed by the study of synaptic connections of identified neurones. *Trends Neurosci.* **13**, 259-265 (1990).
16. Groves, P.M., Linder, J.C. & Young, S.J. 5-Hydroxydopamine-labeled dopaminergic axons: Three dimensional reconstructions of axons, synapses, and postsynaptic targets in rat neostriatum. *Neuroscience* **58**, 593-604 (1994).
17. Schultz, W., Apicella, P. & Ljungberg, T. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* **13**, 900-913 (1993).
18. Parthasarathy, H.B., Schall, J.D. & Graybiel, A.M. Distributed but convergent ordering of corticostriatal projections: analysis of the frontal eye field and the supplementary eye field in the macaque monkey. *J. Neurosci.* **12**, 4468-4488 (1992).
19. Gerfen, C.R., *et al.* D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science* **250**, 1429-1432 (1990).
20. Calabresi, P., De Murtas, M. & Bernardi, G. The neostriatum beyond the motor function: Experimental and clinical evidence. *Neuroscience* **78**, 39-60 (1997).
21. Hikosaka, O., Sakamoto, M. & Miyashita, N. Effects of caudate nucleus stimulation on substantia nigra cell activity in monkey. *Exp. Brain Res.* **95**, 457-472 (1993).
22. Hikosaka, O. & Wurtz, R.H. Modification of saccadic eye movements by GABA-related substances. II. Effects of muscimol in the monkey substantia nigra pars reticulata. *J. Neurophysiol.* **53**, 292-308 (1985).
23. Kato, M. & Hikosaka, O. in *Age-Related Dopamine-Deficient Disorders* (eds. Segawa, M. & Nomura, Y.) 178-187 (Karger, Basel, 1995).
24. Matsumura, M., Kojima, J., Gardiner, T.W. & Hikosaka, O. Visual and oculomotor functions of monkey subthalamic nucleus. *J. Neurophysiol.* **67**, 1615-1632 (1992).
25. Hikosaka, O. in *The Basal Ganglia IV: New Ideas and Data on Structure and Function* (eds. Percheron, G., McKenzie, J.S. & Feger, J.) 589-596 (Plenum Press, New York, 1994).
26. Graybiel, A.M. Building action repertoires: memory and learning functions of the basal ganglia. *Curr. Opin. Neurobiol.* **5**, 733-741 (1995).
27. Hikosaka, O. & Wurtz, R.H. Visual and oculomotor functions of monkey substantia nigra pars reticulata. III. Memory-contingent

visual and saccade responses. *J. Neurophysiol.* **49**, 1268-1284 (1983).

28. Hikosaka, O., Sakamoto, M. & Usui, S. Functional properties of monkey caudate neurons. I. Activities related to saccadic eye movements. *J. Neurophysiol.* **61**, 780-798 (1989).

29. Aosaki, T., *et al.* Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. *J. Neurosci.* **14**, 3969-3984 (1994).

大脳皮質－大脳基底核回路の運動制御における機能

中原 裕之¹、銅谷 賢治²、彦坂 興秀³

¹理化学研究所 脳総研 脳数理研究チーム 和光市広沢 2-1 〒351-0198 Email:
hiro@brain.riken.go.jp

²CREST, 科学技術事業団 & ATR, 京都府相楽郡精華町光台 2-2

³順天堂大学医学部第一生理研究室、文京区本郷 2-1-1

Summary: 本論文では、計算理論の立場から、大脳皮質－大脳基底核－視床回路の機能の概説を述べる。我々は、実験・理論研究を合わせて行うことによって、この大脳皮質基底核系の系列学習における学習・制御について、「多重表現仮説」という作業仮説を提案した。その骨子は、第一に、大脳皮質基底核系の複数のループ、具体的には前頭前野ループと運動ループが、それぞれ異なる座標系を使って系列を学習し、同時に並行して逐次系列の遂行に寄与するというものである。更に、学習ではドーパミン細胞により与えられる強化信号が使われる一方、運動遂行時の両ループのバランスをとるために前補足運動野が機能するというものである。この仮説により多くの実験結果を統合的に理解できる。逐次系列実験課題である2x5課題での実験結果と計算機シミュレーションの結果を交え具体的に述べる。

§1 イントロダクション

運動の学習・制御に関与する大脳皮質―皮質下回路として、二つの重要な回路がある。一つは小脳の回路であり、もう一つが大脳基底核の回路である。この二つの回路は互いに補完しつつ、運動の学習・制御に貢献する。小脳の系は、個々の運動の学習・制御（in movement）に貢献する一方、大脳基底核の系は、逐次的運動や様々な運動の組み合わせ（‘between movements’）に関わっていると以前から考えられてきた。現在においても、この観点は、粗い見方としては妥当であろう。手続き記憶（procedural memory）やスキルの習熟にこれらの系は重要であり、運動の機能解明においてこれらの系の理解は欠かせない。大脳基底核が、逐次的運動、または複雑な運動の学習・制御に関与することを示す顕著な例としては、パーキンソン病・ハンチントン病の症候が挙げられる。黒質の病変を伴うことの多いパーキンソン病は、筋の固縮・振戦あるいは無動症などが特徴的な症候である。一方、ハンチントン病は、舞踏病とも言われ、速い無目的運動が典型的な症候である。このように、大脳基底核が複雑な運動・逐次運動に関わることを示す証拠がある一方で、基底核がどのようにそれらの機能に貢献しているのか、言い換えれば、基底核でどのような計算が実現されているのか、未だ謎の部分が多い。小脳の回路に比べると、大脳基底核回路の計算論的な機能解明は遅れている。しかし、それと同時に、実験的証拠は加速度的に増加しており、それだけに大脳基底核の計算論的研究は、現在やりがいのある研究分野であると言えよう。

我々は、大脳皮質基底核回路の逐次運動の学習・制御における役割の解明を目指し、彦坂研究室で開発された逐次運動系列課題（‘2x5課題’）¹を用いた実験的研究と、計算機におけるモデルシミュレーションを組み合わせ研究を行った^{2 3 17}。大脳皮質―大脳基底核回路の機能を考えていく上で、以下の二点が特に重要である。一つは、その回路全体の特徴、もう一つは、大脳基底核でのドーパミン細胞の報酬予測的な活動の存在である。以下、まずこの二つの特徴について概観する。次に、2x5課題について手短かに説明する。その上で、大脳皮質基底核回路の機能に関する我々の仮説を述べ、実験結果・シミュレーション結果の比較について述べる。

§2 背景

§2.1 大脳皮質基底核系の回路構造

大脳皮質のほぼ全ての領野が大脳基底核に投射する。大脳基底核の入力核である線条体から、基底核内の回路を通じて、大脳皮質の神経活動の影響が出力核である淡蒼球内節部・黒質網様部に伝えられる。そして、視床を介して、前頭葉に投射が返されている。なお一部は、上丘へも送られている。注目されるのは、この回路系は、いくつかの分離された回路構造をしていることである^{4 5}。例えば、運動ループと呼ばれる系では、運動関連領野から線条体後部（被殻）に投射し、そこから基底核の出力核→視床→補足運動野及び運動前野・第一次運動野へと返されている。一方、前頭前野ループと呼ばれる系では、前頭前野および一部の頭頂野からの投射は、線条体前部（尾状核）に到達し、前頭前野に投射が戻るような回路を形成する。他にもいくつかの系があり、これらの分離された系を束ねたような回路を大脳皮質・大脳基底核は形成している。この回路構造は何のためにあるのかという疑問が当然湧いてくるわけだが、それについては後に述べる。

§ 2.2 ドーパミン細胞による報酬の予測誤差信号

大脳基底核内部の構造の特徴としては、ドーパミン細胞の線条体に及ぼす影響が非常に強いことが挙げられる。ドーパミン細胞が、食物などの様々な‘快’を与えるもの（総じて報酬信号と呼ばれる）に反応することは古くから良く知られていた。近年の研究は、更に、ドーパミン細胞の神経活動が報酬信号の予測誤差（‘強化信号’）をコードしていることを示唆している⁶。機械学習の分野で研究されてきた‘TD学習’と呼ばれる学習方式では、得られた報酬そのものではなく、将来予測される報酬の時間的変化（Temporal Difference）をもとに行動学習が行われる。ドーパミン細胞の活動は、この報酬予測のTD信号の振舞いに非常に近い⁷。紙面の制約より、ここではTD誤差について解説しない^{8,9}。興味深いのは、冒頭に述べたように大脳基底核が逐次運動の学習・制御に貢献するという実験的・臨床的結果が得られている一方で、ドーパミン細胞の反応特性が逐次運動の学習に適しているという知見が独立に得られたことである。

§ 2.3 2x5課題：逐次運動系列課題

我々人間の日常は、数多くのスキル・手続き記憶に支えられている。歯を磨く・洋服を着る・ハシを使うなど、特別に意識せず我々が行う日常の活動も、元を正せば、幼少時から時間をかけて獲得されたスキルである。これらの無数の習熟したスキルによって普段の生活が営まれる一方で、我々は、例えば、ある日からテニスを学び始める、ワープロの使い方を学ぶなど、新しいスキルを学び始める。このように、我々は、生涯、絶えず新しい様々な課題をこなし、時には長い時間をかけて新しいスキルを獲得し、習熟したスキルについてはほとんど無意識に遂行できるようになる。これらは手続き記憶の重要な特徴である。これらの観察を念頭に、2x5課題は開発され（図1 A）¹、サル及び人間の実験が行われた。2x5課題では、「ハイパーセット」と呼ばれる五つの「セット」を一定の順番で並べた逐次系列が一つの実験単位になっている。個々のセットでは、16のLEDボタンのうち二つが点灯し、被験者は、ある決まった順番でそのボタンを押さなければならない。一つのセットで間違えることなく押せると、次セットの二つのボタンが点灯する。個々のセットでのボタン押しの正しい順番は試行錯誤しながら見つけなければならない。五つのセット（計：2x5=10のボタン）を間違えることなく押せると、ハイパーセットが一試行成功したことになる（このときに、サルは最大の報酬が得られる）。1ブロックの実験では、20回の成功試行をするまで、同じハイパーセットが続けて提示される。サルは、一日に約10~20個のハイパーセットを経験するが、その約半数は、その場限りで経験するだけだが、残りのハイパーセットは毎日同じものを経験する。この後者のハイパーセットに対する成績は、何日もの経験の後に上昇していった（図1 B）。実験では、これらのトレーニング期間中と、更にトレーニング後に、様々な実験操作を加えることによって、サル（あるいは人間）のパフォーマンスを記録した。例えば、一つのハイパーセットを終了するまでに要した試行数は、パフォーマンスの一つの目安である。また、脳の各部位の神経細胞活動記録（人間の場合は、fMRI^{10,11}）が調べられた。

§ 3 系列運動の学習・制御のための大脳皮質基底核回路のメカニズム

まず手短かに、系列運動の学習・制御を実現する大脳皮質基底核回路のメカニズムについて

て我々の提案する作業仮説の骨子を述べておこう（図2）。i) 皮質-基底核回路の異なるループは異なる座標系を用いている。特に、前頭前野ループ・運動ループは、それぞれ視覚座標系・運動座標系を用いている（図1 C）。これは、各々のループに関わる領野の入出力の解剖学的特徴と神経細胞の反応特性に基づいて与えられた仮説である。以後、前頭前野ループを視覚ループと呼び、議論を簡明にするために、大脳皮質基底核のループのうち、視覚・運動ループのみを議論する；ii) 異なった座標系を持つが故に、視覚・運動ループに異なる特性が現れる。視覚ループは学習が速い一方、運動ループは学習が遅いが一旦学習すれば速い運動遂行が可能である（図2）；iii) ドーパミン細胞から得られる強化信号を利用して、視覚・運動ループは、同じ運動系列を並列的に（但し学習速度は違うが）学習する（図2）；iv) 視覚・運動ループは並列的に学習する一方で、運動の遂行にも並列的に関与する。この時、当然、二つのループが異なる遂行命令を出す状況が生じうるが、そのような時に両者の出力のバランスをとって最終的な遂行命令を出す必要がある。その時の調整を行う機能を、ここではコーディネーターと呼ぶ。我々は、コーディネーター機能（あるいは少なくともその一部）は、前補足運動野に存在すると考えている（図2, 3）；v) 視覚ループの出力命令は視覚座標系を用いて与えられることになるが、実際に腕などを動かすためには運動座標系に変換される必要がある。この機能をトランスレーターと呼ぶ。トランスレーターが存在すると思われるのは運動前野腹側部である（図3）。

我々は、上記の作業仮説を、様々な実験結果・理論的考察・計算機モデルのシミュレーションに基づいて提案した^{2 3 12 17}。実際、構築した計算機モデルで2x5課題のシミュレーションを行うと、実験結果と広範な範囲で良い対応が得られている。以下、その対応を説明すると同時に、上記の提案について概説する。同時に、以下説明される諸点が、手続き記憶の重要な特徴に対応していることにも注意して欲しい。

§3.1 複数の学習段階

シミュレーションでは、実験同様まずトレーニング期間が設けられた。これによって、モデル系はいくつかの系列に習熟した。シミュレーションでは、実験で見られた三つの異なる学習段階のうち、二つの段階が現れた。その場限りで経験する逐次系列（ハイパーセット）が20試行成功するまでのパフォーマンスを調べると、最初誤りが多く、後半の方になればなるほど成功する試行が続けて現れることが観察された（第一のレベル）。その一方で、ある逐次系列を何日も経験するうちに、20試行成功するまでの総試行数は徐々に減っていく（第二のレベル）。我々のモデルでは、この第一のレベルでは、主として視覚ループが系列の遂行を行っており、運動ループがその系列に習熟していくにつれ、第二のレベルが現れてくる¹。

§3.2 学習の転移

野球好きの人でも、右投げの人が左手で投げるのはどこかきこちない。手続き記憶（運動記憶）は、その運動部位に、一定程度依存して学習されるからである。実際、2x5課題で右手で習熟した系列を左手で、あるいは左手で習熟した系列を右手で行わせると、成績が悪くなる。その一方、全く新しい系列を行うよりは、少ない試行数で終わることができる^{13 14}。我々のモデルでは、反対側の手を使うときには、運動ループの記憶を使用することができなくなるので、成績は下がるが、その一方、視覚ループに蓄えられた記憶を使用することができるので、新しい系列よりは成績が良くなると考えられる。

§ 3.3 系列に依存した記憶

手続き記憶の重要な特性として、その記憶が系列に依存して蓄えられていることが挙げられる。実際、2x5課題でも、ある習熟された系列（ハイパーセット）をとりあげ、その系列内に提示するセットの順序を変えると、その成績は、新たな系列遂行時の成績とほとんど変わらない³。これは、我々のモデルでは、記憶が系列依存的に蓄えられることに対応している。

§ 3.4 視覚ループと運動ループの機能の分化

サル¹⁵の2x5課題の実験で、十分なトレーニングの後に、脳の各部位にムシモール（GABA agonist）を注入するという実験を行った。線条体前部と後部に別々に注入したときには、両者の結果を比較すると、前部注入は新しい系列の成績を悪くし、後部注入は習熟した系列の成績を悪くするという機能の分化が観察された¹⁶。モデルでも同様の現象が見られた。

§ 3.5 コーディネーターとしての前補足運動野

前補足運動野のムシモールの注入実験では、新しい系列のパフォーマンスのみ悪くなるのが観察されたが、モデルでも同様の現象が見られた¹⁵。コーディネーターの基本的な役割は、新しい系列の遂行時には、運動ループでの学習がまだあまりできていないので、主に視覚ループに依存して出力命令が決定されるようにすることにある。

§ 3.6 ドーパミン欠損のシミュレーション

モデルのシミュレーションで、トレーニングが終わった後で、ドーパミン細胞が線条体前部あるいは後部で欠損した場合を試した。すると、前部で欠損があった時のみに、新しい系列に対するパフォーマンスが低下した。この結果は、視覚ループの早期の学習の重要性を示す一方、ドーパミン細胞から与えられる強化信号が学習に重要であることも示している。なお、この実験は行われていない。

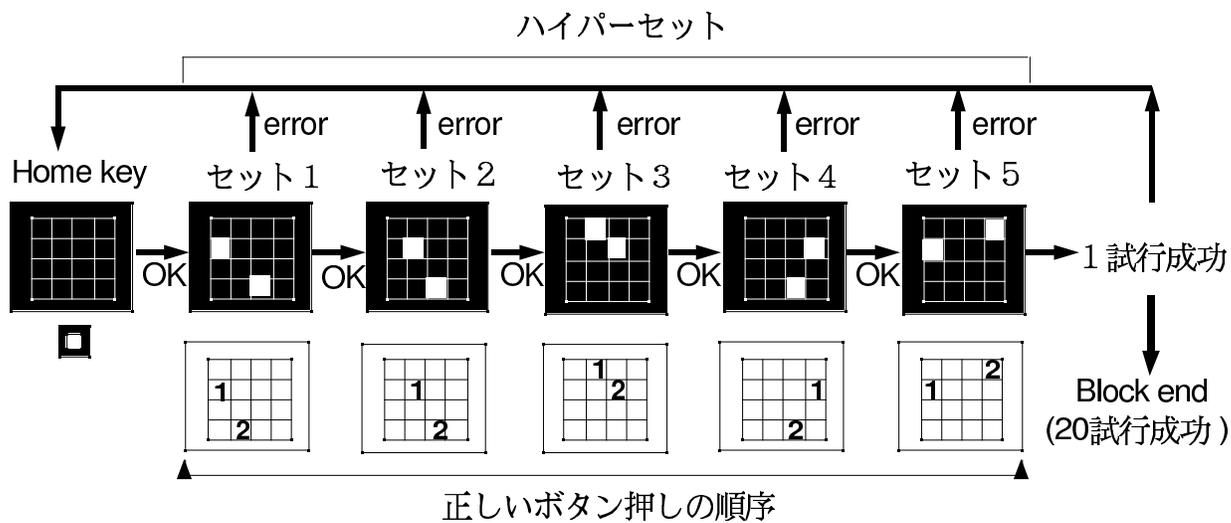
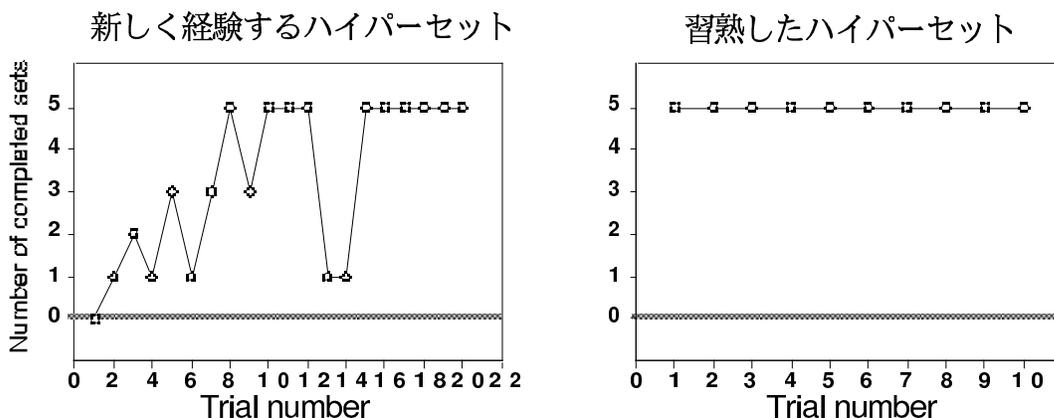
§ 4 結語

ここでは、大脳皮質基底核回路の系列運動の学習・制御に関する我々の作業仮説について概説した。この仮説は、まだ謎の多い大脳基底核系の計算論的解明の出発点の一つとして有望であると我々は考えている。同時に未だ詰められていない点も多い。例えば、大脳皮質での補足運動野でのムシモール注入では¹⁵、その結果は我々のモデルと必ずしも全て一致していない^{15 12}。また、帯状皮質運動野の報酬予測に関連するような神経活動と前補足運動野のコーディネーターの機能の関係も今後の課題である。これらの点も含め、今後も検討を続けていきたい。

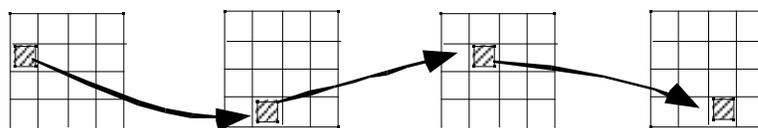
1. Hikosaka, O., Rand, M.K., Miyachi, S. & Miyashita, K. Learning

of sequential movements in the monkey - Process of learning and retention of memory. *Journal of Neurophysiology* **74**, 1652-1661 (1995).

2. Nakahara, H. Sequential Decision Making in Biological Systems, Ph.D. Thesis, University of Tokyo (1997).
3. Hikosaka, O., et al. Parallel neural networks for learning sequential procedures. *Trends in Neurosciences* **22**, 464-471 (1999).
4. Alexander, G.E. & Crutcher, M.D. *Functional architecture of basal ganglia circuits: neural substrates of parallel processing* 1-266-271 (1990).
5. Hoover, J.E. & Strick, P.L. Multiple output channels in the basal ganglia. *Science* **259**, 819-821 (1993).
6. Schultz, W., Apicella, P. & Ljungberg, T. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience* **13**, 900-913 (1993).
7. Schultz, W., Dayan, P. & Montague, P.R. A neural substrate of prediction and reward. *Science* **275**, 1593-1599 (1997).
8. Barto, A.G. in *Models of information processing in the basal ganglia* (eds. Houk, J.C., Davis, J.L. & Beiser, D.G.) 215-232 (MIT Press, Cambridge, MA, 1995).
9. Houk, J.C., Adams, J.L. & Barto, A. in *Models of information processing in the basal ganglia* (eds. Houk, J.C., Davis, J.L. & Beiser, D.G.) 249-270 (MIT Press, Cambridge, MA, 1995).
10. Hikosaka, O., et al. Activation of human presupplementary motor area in learning of sequential procedures: A functional MRI study. *Journal of Neurophysiology* **76**, 617-621 (1996).
11. Sakai, K., et al. Transition of brain activation from frontal to parietal areas in visuo-motor sequence learning. *Journal of Neuroscience* **18**, 1827-1840 (1998).
12. Hikosaka, O., et al. in *The New Cognitive Neurosciences* (eds. Gazzaniga, M.S.) 553-572 (The MIT Press, Cambridge, MA, 1999).
13. Rand, M.K., Hikosaka, O., Miyachi, S., Lu, X. & Miyashita, K. Characteristics of a long-term procedural skill in the monkey. *Experimental Brain Research* **118**, 293-297 (1998).
14. Rand, M.K., et al. Characteristics of sequential movements during early learning period in monkeys. *Experimental Brain Research* **131**, 293-304 (2000).
15. Nakamura, K., Sakai, K. & Hikosaka, O. Effects of local inactivation of monkey medial frontal cortex in learning of sequential procedures. *Journal of Neurophysiology* **82**, 1063-1068 (1999).
16. Miyachi, S., et al. Differential roles of monkey striatum in learning of sequential hand movement. *Experimental Brain Research* **115**, 1-5 (1997).
17. Nakahara, H., Doya, K., Hikosaka, O. & Nagano, S. Multiple representations in the basal ganglia loops for acquisition and execution of sequential motor control. *Society for Neuroscience Abstracts* **23**, 778 (1997).

A**B****C**

視覚座標で表示された系列



運動座標で表示された系列

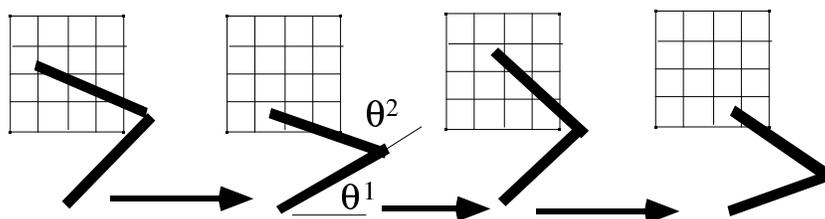


図 1 : 2x5 課題と視覚・運動座標で表示された系列.

A. 2x5課題 B. 新しいハイパーセットと習熟したもののパフォーマンス例 C. 視覚・運動座標で表示された系列

視覚ループと運動ループの
コーディネーション

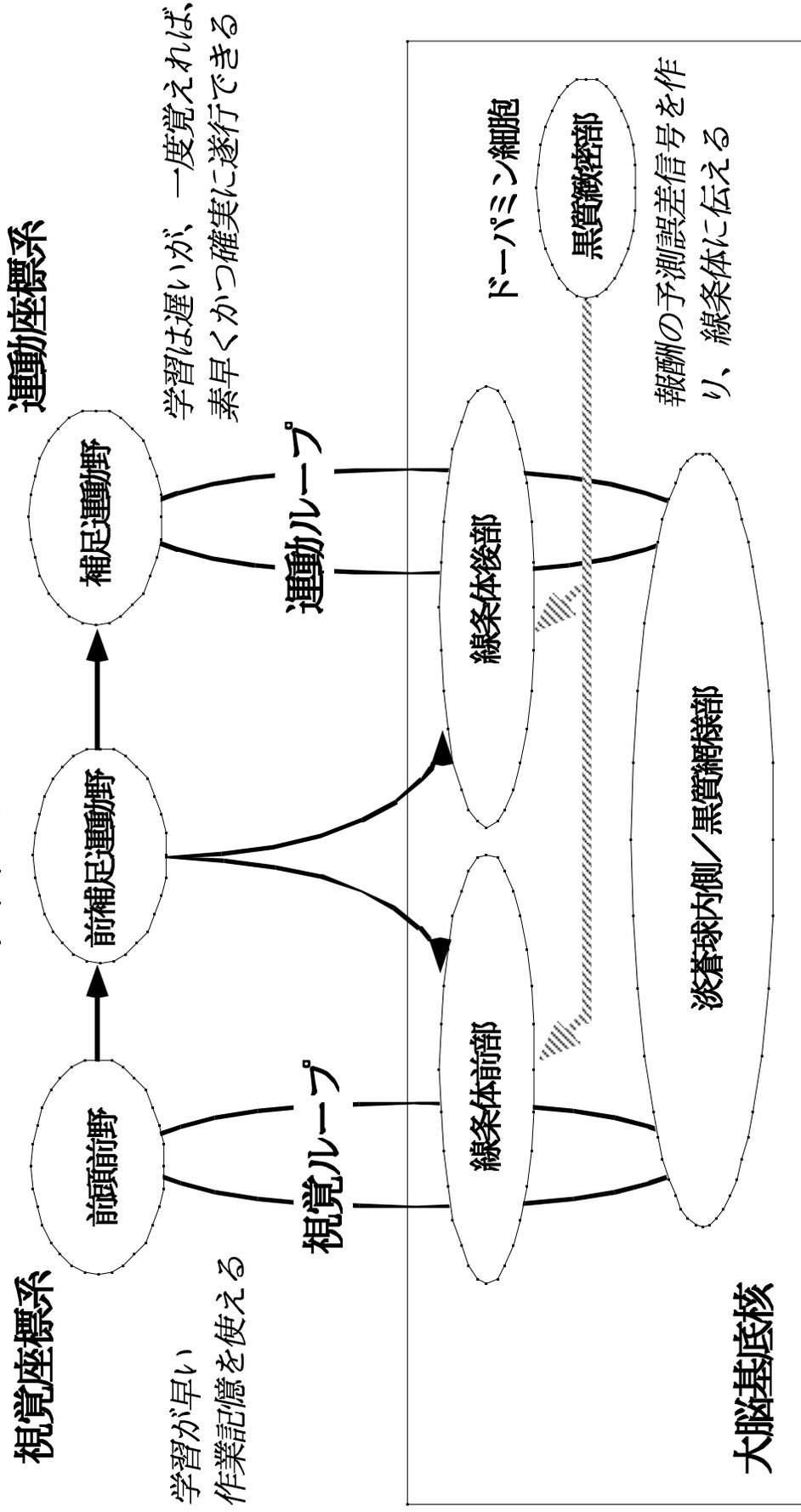


図2：逐次運動の学習・制御のための大脳基底核ループの機能

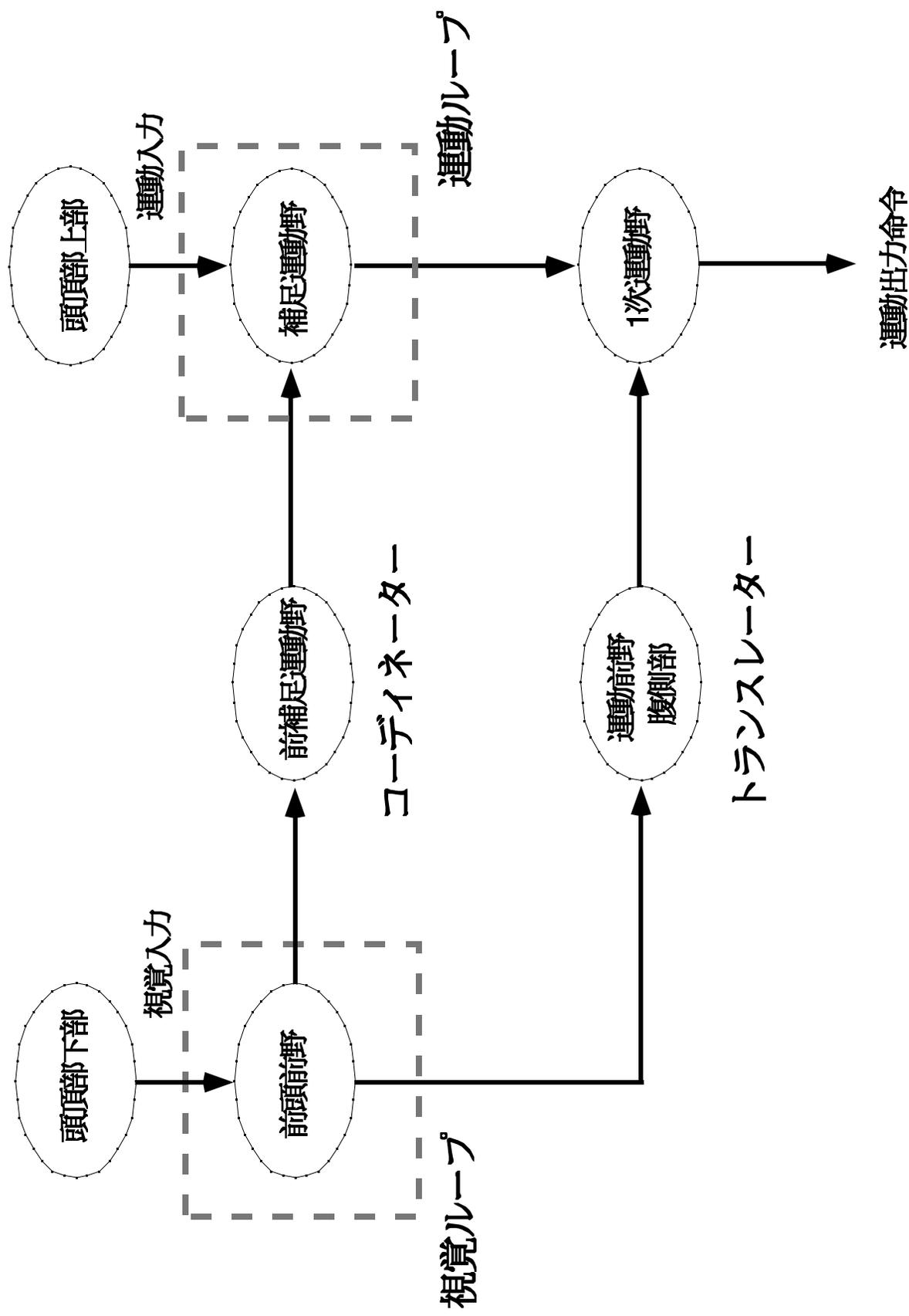


図3 逐次運動の学習・制御における皮質領野の機能
 (大脳基底核とのループの部分は省略して表示してある)

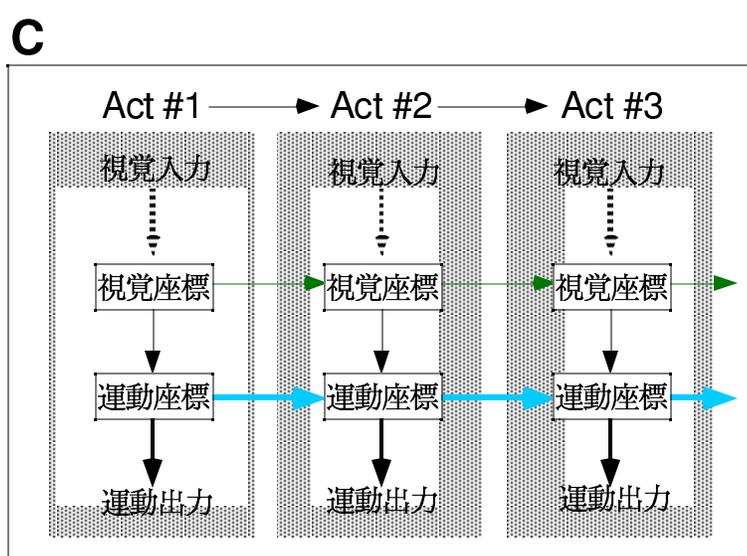
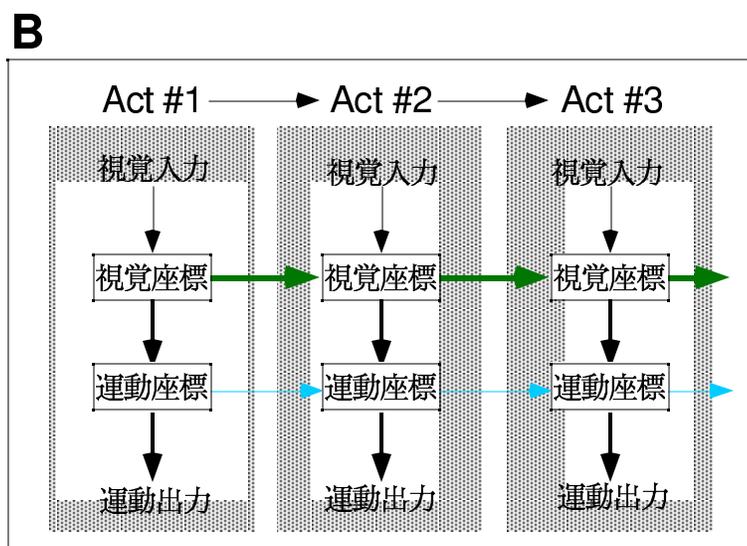
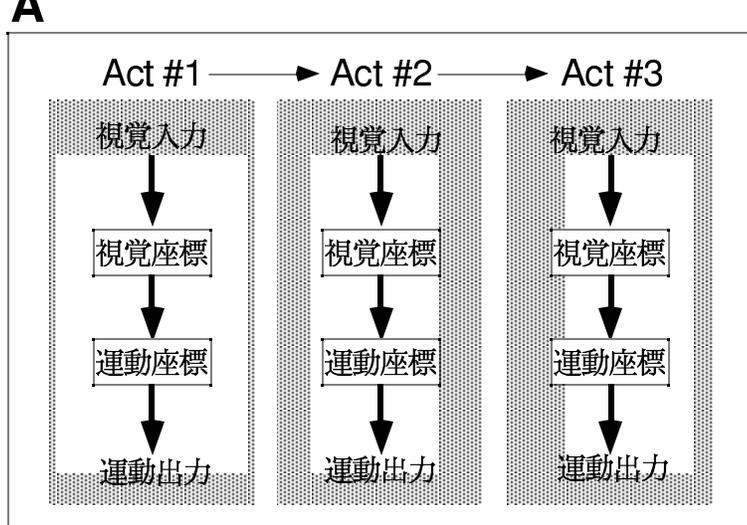


図4：視覚・運動座標を用いた逐次系列の遂行。

学習の最初的时候には、各視覚入力と運動出力の組み合わせを学習する (A)。学習が進むと、系列のパターンそのものをコードできるようになるが、視覚座標での学習が先に現れ (B)、十分に学習が進むと運動座標でコードできるようになる (C)。この段階では、視覚入力にさほど頼らずに、記憶された運動の系列パターンに基づいて遂行が可能になる。

4 日目 統計的予測 大脳皮質

コーディネータ：柏野牧夫，岡田真人

1 . 講義概要

大脳皮質における感覚情報処理は，並列性と階層性という構造的特徴を持っている．また外界を内部モデルとして取り込むことで外界に対して適応的に変化するという機能的特徴を持っている．この二つ特徴を軸に，主として計算理論と心理物理の立場から，大脳皮質の処理の本質を考える． Rao 先生は，計算理論の立場から，知覚の過程は受動的に入力情報を分析するだけではなく，大脳皮質内の内部モデルによって能動的に入力の予測を試みる過程であると考える．そして，その二つの過程が，階層的な構造の中で，ボトムアップとトップダウンの情報の流れとして実現されていると主張する． Watanabe 先生もボトムアップとトップダウンの情報の流れに注目する．視覚運動情報処理は局所的な運動検出に始まり，高次レベルで大域的な運動情報が処理される．このボトムアップの階層性に対して，トップダウンの注意による修飾効果は注意特徴に応じて選択的に変化する．また，外界への適応のための神経の可塑性を示す知覚学習にも階層構造が関係していることを説明する．

知覚学習と並んで，感覚系が外界に対して適応的に変化していることを示す現象が順応である．西田は視覚における順応現象の代表例である運動残効を切り口に，視覚運動情報処理の並列・階層構造を明らかにし，人間の脳が運動情報に基づいて未来の位置を予測していることを示す．柏野は，聴覚の音源定位に関する順応現象を取り上げ，音源定位機構を解析するとともに，順応や知覚学習が音源弁別能力を向上させるメカニズムについて考える．そして，順応の背後にある神経の可塑性によって大脳皮質は外界を内部モデルとして取り込み最適に入力情報を符号化しているという Barlow の仮説について論じる．

2 . 必読文献

- [1] R. P. N. Rao and D. H. Ballard (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, vol.2, no.1, 79-87.
- [2] Watanabe, T. & Miyauchi, S. (1998). Interactions in visual motion processing: Psychophysical and brain imaging studies. In T. Watanabe (Ed.), *High-level motion processing -- Computational, biophysiological and psychological approaches* (.pp. 95-114), MIT Press.
- [3] Barlow, H. B. (1990). A theory about the functional role and synaptic mechanism of visual aftereffects. In C. Blakemore (Ed.), *Vision: coding and efficiency* (pp. 363-375). Cambridge, UK: Cambridge University Press.

Predictive Coding in Recurrent Neocortical Circuits

Rajesh P.N. Rao

Salk Institute for Biological Studies, USA

Neocortical circuits are dominated by recurrent feedback connections. If cortical area A projects to cortical area B, then area B invariably sends feedback connections to area A. Similarly, within a given cortical area, there exists massive recurrent excitatory feedback between pyramidal neurons due to local horizontal connections. What is the role of feedback in cortical information processing? I will review recent models that postulate that (a) feedback connections between cortical areas instantiate statistical generative models of cortical inputs, and (b) recurrent feedback connections within a cortical area encode the temporal dynamics associated with these generative models. The resulting network allows predicting coding of spatiotemporal inputs and suggests functional interpretations of nonclassical surround effects in the visual cortex on the basis of natural image statistics. We show that recent results on spike timing dependent plasticity in recurrent cortical synapses are consistent with such a model of cortical feedback and present comparisons of model simulations to electrophysiological data from awake monkey visual cortex.

Roles of attention and form in visual motion processing: Psychophysical and brain imaging studies

Takeo Watanabe

Department of Psychology, Center for the Adaptive System, Boston University

INTRODUCTION

It goes without saying that the human visual information processing system is very complex. Marr (1982) proposed that several principles should be taken into consideration to improve understanding of this complex system. These include the *modularity principle* and the *feedforward principle*. The modularity principle assumes that the visual system consists of several processes that are relatively independent of each other in early stages of visual processing (front-end processing). Therefore, each of these processes can be largely examined independently without considering the others. The feedforward principle assumes that the majority of visual functions can be understood without taking feedback effects into consideration.

While research that follows these principles has greatly contributed to understanding some important aspects of visual information processing, other important aspects of information processing have been relatively neglected, especially, interactions between different processes and feedback from higher to lower levels (however, see Grossberg & Mingolla, 1985; Koch & Davis, 1994).

Research on motion perception is not exceptional in this sense. Over the last decade, as in other aspects of visual processing, the majority of research on the processing of velocity of a moving object has mainly explored feedforward processing within the motion module (for a review, Hildreth & Koch, 1987; Nakayama, 1985; Snowden, 1992).

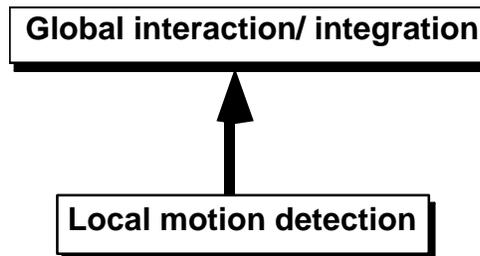


Figure 1 Two stage model

One of the most successful models made from this viewpoint is the *two-stage model*. As shown in Figure 1, this model assumes that visual motion processing includes two stages¹ of analysis (Adelson & Movshon, 1982; Hildreth & Koch, 1984; Movshon, Adelson, Gizzi & Newsome, 1986; Snowden et al, 1991; Welch, 1989). The first stage is supposed to measure *local components of motion*. The directions of locally measured motion components of an object have been reported to be orthogonal to the object's local contours (e.g., Gizzi, Katz, Shumer and Movshon, 1990; Movshon et al, 1986). These locally measured motion components, therefore, do not necessarily correspond to the perceived motion of the whole object. In the second stage, the locally measured motion signals for a single object are integrated. For example, as shown in Figure 2, when two moving gratings overlap, instead of moving in the two different directions orthogonal to the two gratings (Figure 2 (a, b)), they appear to make a "plaid", moving in one motion direction (Figure 2 (c)) (Adelson & Movshon, 1982; Movshon et al, 1986)².

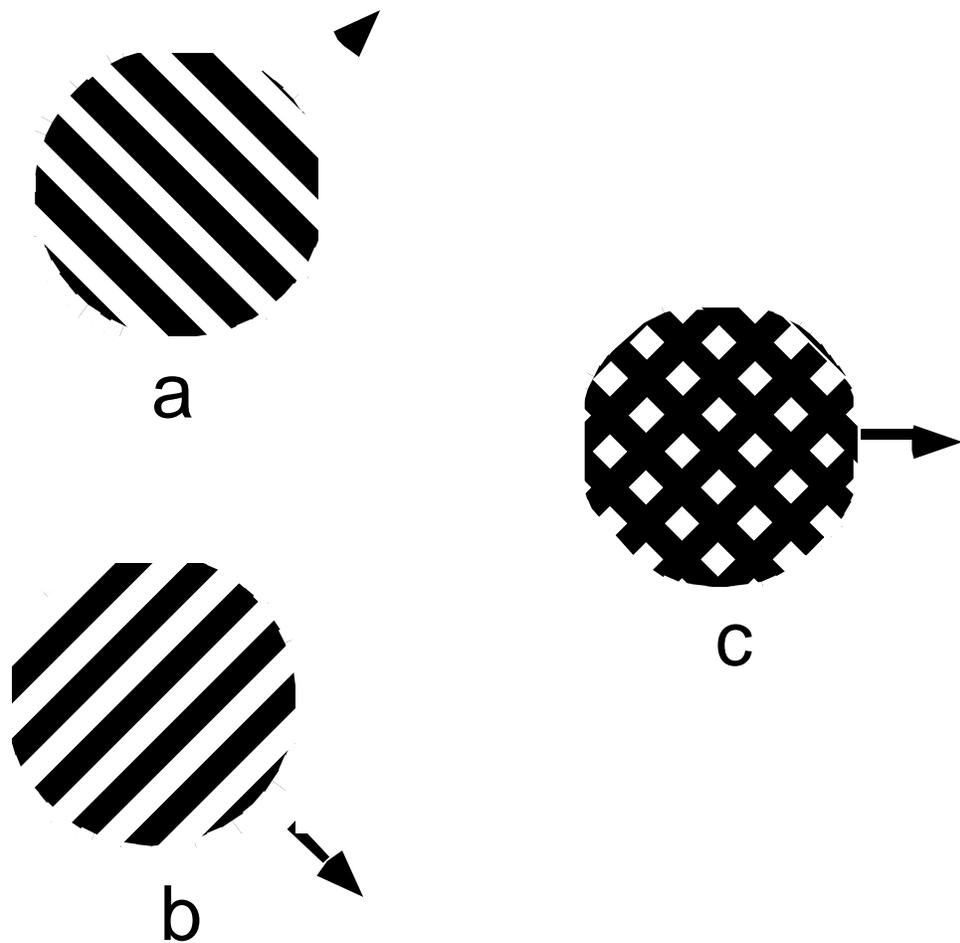


Figure 2: The gratings appear to move in the directions orthogonal to the stripes as in a and b. However, when these two gratings are superimposed, they appear to make a "plaid", moving in one motion direction (Adelson & Movshon, 1982).

The first stage seems to mainly involve V1 and to a certain extent MT, while the second stage of processing may occur at MT, but not at V1. Movshon et al (1986) found that when a plaid consisting of two moving gratings is presented, they found cells in V1 tuned to directions orthogonal to the gratings. However, no cell in V1 was found to be tuned to the integrated motion direction of the plaid. On the other hand, 40% of the cells in MT were found to be tuned to the motion directions orthogonal to the gratings, and 25% of the cells to the integrated motion direction (also see Rodman & Albright, 1987).

Snowden et al have also proposed the two stage model. Using a "transparent" stimulus composed of two sets of random dots moving in opposite directions, it has been reported that cells at V1 tuned to each of the two opposite motion directions gave a similar response to the transparent stimulus and the single surface stimulus composed only of the dots moving in one of the two directions. On the other hand, the activity of the cells in MT stimulated by the transparent stimulus was found to be constantly suppressed³. That is, local velocities are detected in V1 and the signals of these velocities are interacted in MT (Snowden, Treue, Erickson and Andersen, 1991).

Although this two-stage model has made a great contribution to understanding early motion processing, it considers only feedforward processing within the motion module. In order to better understand motion processing, it should be examined from a viewpoint of visual processing as an interaction of interdependent processes. Recent physiological and anatomical studies show that there are massive reciprocal projections (De Yoe & Van Essen, 1988). These projections may be used in at least two ways in motion processing. The first is for carrying signals from higher-level to lower-level motion processing (feedback projections). When a subject is instructed to pay attention to a whole or part of a moving object, this kind of active attention should originate in a high-level stage of neural information processing (e.g., Corbetta et al, 1991; Posner & Corbetta, 1992) and may influence low-level visual processing. The second way is for reciprocal projections between motion and other types of information processing (inter-module projections). A signal which is sent from motion processing to another processing and that is modified there may be sent back to the motion module in order for the visual system to produce a unified surface representation (Cavanagh, 1987; He & Nakayama, 1992) or representations that are consistent with each other. These two possible kinds of information—feedback projections and inter-module projections—flows should be considered in relation to the two-stage model.

Here, we will discuss our research on early motion processing in relation to both attention (feedback projections) and form processing (inter-module projections)

EFFECT OF ATTENTION ON EARLY MOTION PROCESSING —FEEDBACK PROJECTIONS—

There has been a good deal of controversy about whether or not attention influences very early stages of visual information processing. A considerable amount of work has suggested that attention influences relatively high-level processing, but not the feature analysis level. Treisman and her colleagues built a feature integration theory in which one role of attention is to integrate visual features that are independently processed at lower level stages (Treisman & Gelade, 1981). Neurobiologically, Moran & Desimone (1985) have found that the response of cells in areas V4 and the inferior temporal area (IT) of macaque monkeys to an unattended stimulus is dramatically reduced. However, they did not find such response reduction in cells in V1, where local measurement in various stimulus dimensions occurs (Hubel & Wiesel, 1959; 1962). On the other hand, Motter (1993) found that the activity of orientation-tuned cells at V1 was enhanced when subjects (macaque monkeys) attended to a specific orientation.

With regard to the effect of attention on motion processing, Cavanagh (1991) found that motion can be driven by attention in the absence of low-level motion signals⁴. Culham and Cavanagh (1994) also found that motion capture is influenced by attention. These findings suggest that attention is quite influential in motion processing. Recently, it has been found attention modulates MT and MST in monkey by means of electrophysiology (Treue & Maunsell, 1996) and the human homologues of MT and MST by f-MRI technique (O'Craven & Savoy, 1995). However, it is not clear whether attention influences each of the first stage of motion processing where local component motion directions are measured and how attention influences the first and second stages.

In order to clarify these questions, the roles of attention in the first, local motion component measurement stage and in the second, integration stage should be examined. To accomplish this, several questions must be addressed: Can the visual system selectively attend to a locally-measured component motion? If so, does attention directly activate the local mechanism or unit for the component motion? Does attention alter the interactions of the signals from locally measured-component motions? Does attention operate on different stages of motion processing, depending on the nature of the task?

If these questions about the influence of attention on motion processing are clearly answered, we will attain a better insight not only into the role of attention in motion processing but also attention in general.

The role of attention in the first stage with f-MRI

The first question is whether attention can modulate component motion directions that are measured in the first stage. Assume that a complex object is shown moving at a velocity that is determined by the integration of locally measured component motions. If directing attention to a local contour of the complex object makes us perceive it moving in the direction orthogonal to the local contour, we may say that attention modulates the component motion.

The previous studies have used either moving random dots or a plaid pattern. Both stimuli contain problems to be used for the purpose of examining the effect of attention on local component motion.

Lankheet and Verstraten (1995) found that when the subject was asked to direct attention to one of two surfaces of random dots moving in opposite directions, the adaptation to the attended direction was higher. Although their finding itself is interesting and very important, using random dots causes a problem for our experiment in that it is not clear whether the subject directed attention to a local motion or a spatially integrated surface moving in a particular direction.

The plaid pattern is also problematic for the purpose of our study. It has been pointed out that the intersections of two gratings of a plaid pattern, which move in the same direction as the perceived motion direction of the plaid as a whole, may be used as a local tracking cue to the plaid motion (Gorea. & Lorenceau, 1991; Alais, Wenderoth, & Burke, 1994). Alais et al (1994) found that the inter-ocular transfer of the motion aftereffect to the integrated motion direction of a plaid pattern was not perfect, suggesting that the monocular motion unit for the motion direction identical to the integrated motion direction is actually involved when the integrated motion is perceived. This finding is in accordance with the view that intersections are used as a local tracking cue. Thus, the plaid motion may not be entirely determined by the internal integration of locally measured component motions.



Figure 3 : The test stimulus consisting of the non-parallel stripes moving either to the left or to the right in a circular aperture.

In order to get around these potential problems to examine the effect of attention to a local component motion, Watanabe (1995) have developed a new stimulus that is shown in Figure 3. It consisted of a sequence of black wedges moving against a white background toward the right or the left in a circular aperture. The black wedges (0.5 cd/m^2 in luminance) were 30 deg in angular subtense and moved either to the right or left at a speed of 2.5 deg/sec within a white circular aperture (32.0 cd/m^2) against a dark gray background (16.0 cd/m^2). The radius of the circular aperture was 7.5 deg . A light gray cross (24.5 cd/m^2) was presented in the center of the circle as a

fixation point. The motion of the wedges as a whole should be the result of the integration of the signals from component motions that are measured along the right and left contours of the stripes. Both the IOC hypothesis² and the vector summation hypothesis predict that the motion direction(s) of the wedges as a whole should be either rightward or leftward as represented by the arrow in Figure 3. In this new display, although indeed there are unambiguously moving features, namely the corners near the top or bottom of the aperture, they do not move in the same directions as the integrated motion direction. Thus, when the wedges are perceived to move in the integrated motion direction (either rightward or leftward), unlike the intersections in the plaid, these features should not be tracked as a local cue for the motion direction of the plaid as a whole.

When these wedges are observed without allocating attention to any particular place or any particular motion while fixating the fixation point in the center of the display, the wedges appear to move to the left or right in the true direction of motion. However, once attention is directed selectively at one of the wedge contours, the whole wedges appear to move in the direction orthogonal to the attended contour. In the experiment, after the disappearance of the wedge stimulus, the subject was instructed to adjust an arrow to the perceived motion direction of the wedge. The experimental result with 6 subjects (Figure 5) confirmed the observation. Eye movements were observed on a monitor [Panasonic wv-5410] connected to a camera [Panasonic wv-1850] directed at the right eye of the observer. No significant pursuit eye movements were observed throughout the experiments. Notice that the component motion direction detected at a contour is believed to be orthogonal to the contour (Movshon et al, 1986). Thus, these results suggest that attention modulates a local motion direction.

Next, we presented the subject with the test stimulus shown in Figure 2 and examined blood flow change (activation) by means of functional Magnetic Resonance Imaging (f-MRI) technique. So far we found that for all 6 subjects V1 was more activated when they directed attention at one of the contours of the wedge than when they observed the test stimulus passively (Miyachi et al, 1995, 1996). This is shown in the functional map in Figure 4.

Such an increase in activation in V1 in the attention condition was not found with any subject when the same wedge was presented stationary. This result suggests that the V1 activation is due to attention to a local motion rather than attention to a spatial location.

In another control experiment, we presented a black rectangle moving in the circular aperture with the same procedure as in the first experiment. The horizontal width of the rectangle was the same as the average width of the wedge in Figure 2. In that case, no significant activation difference was found in V1 with any of the six subjects between the attention condition and the passive condition. Since the orientation of both side of the rectangle was vertical, the direction of the local motion should be the same as the direction of the integrated motion. Thus attention might not have to strongly activate the unit for the local motion direction. Similar results were reported with a monkey brain (Motter, 1991).

This finding shows that attention influences V1 where local motion directions are first measured (Movshon et al, 1986). How is V1 influenced by attention? One possibility is that the unit (in V1) which is responsible for a local motion was selectively more activated in the attention tasks. However, the mere fact of an activation increase in V1 may not filter out another possibility that this activation increase might be merely due to the general arousal of V1, for example, because of the difficulty of the attention task as compared with no attention task (see Wurtz et al, 1980).

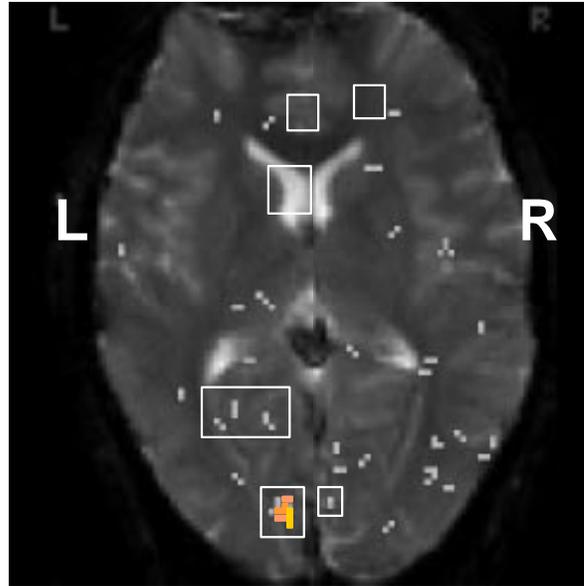


Figure 4 : Functional map of one observer (SM) overlaid on the anatomical map for the right side (right image) attention conditions after the no attention condition results have been subtracted.

In order to examine which possibility is more likely, Watanabe (1995) conducted psychophysical experiments that measured motion aftereffects. It is known that after we perceive an object moving in a direction for several minutes, a physically stationary object appears to move in the opposite direction to that of the previously viewed moving object. This motion aftereffect is usually thought to be due to the adaptation of a set of units tuned to the direction of a moving object. This adaptation in turn causes an imbalance between outputs from the two sets of units responsible for opposing motion directions (e.g., Anstis, 1986). Motion aftereffects have been regarded as a strong and objective measure used to infer the kind of inner motion process activated while the moving object was observed before the occurrence of the aftereffect (e.g., Alais, Wenderoth, & Burke, 1994). Several researchers have pointed out that attention can alter strength and direction of motion aftereffect (Chaudhuri, 1990; Lankheet & Verstraten, 1995; Verstraten et al, 1994). This suggest that attention influences motion processing. However, it was not clear which stage of motion processing is influenced by attention.

Watanabe (1995) compared the strength of motion aftereffects with inter-ocular viewing to that with monocular viewing, in both attention and non-attention conditions. On each trial, the same wedge stimulus as used for the f-MRI experiment (Figure 3) was presented as an adaptation stimulus for 1 min and was followed by a test stimulus consisting of black and white parallel stripes whose orientation was varied from trial to trial. The adaptation and test stimuli were presented either with monocular (same eye) or inter-ocular (one eye to another) viewing . As in the previous experiment, there were two attention conditions and one non-attention condition. The duration of the motion aftereffect was measured as an index of its strength. Before the experiment, each subject conducted approximately 100 practice trials to reduce response variability.

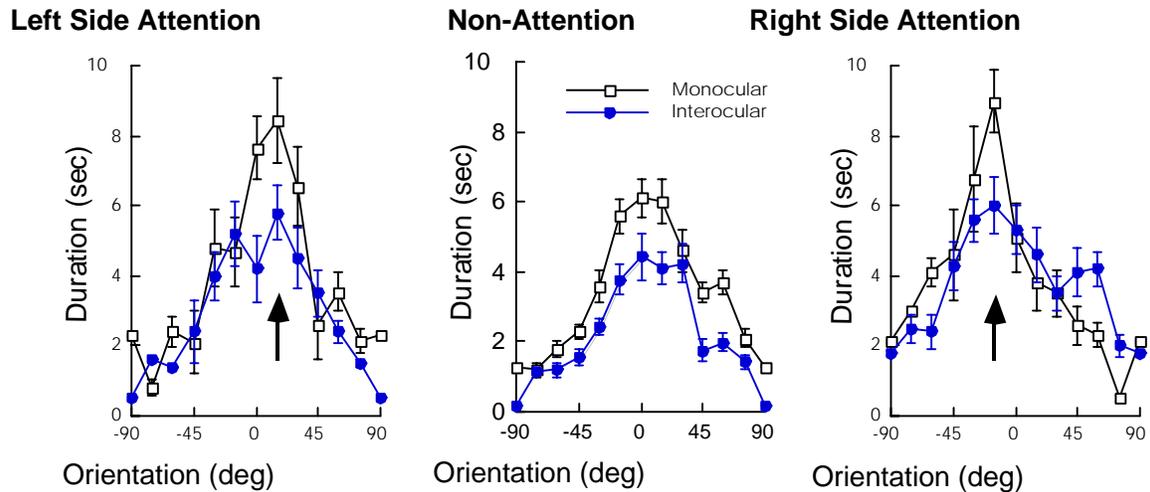


Figure 5: The duration of motion aftereffect as a function of orientation of the test stripes for the left side attention (left), the non-attention (center), and the right side attention (right) conditions. 0 and 90 (-90) deg represent vertical and horizontal orientations. The arrows represent the orientation of the side where attention was directed.

Figure 5 shows the duration of motion aftereffect as a function of test stripe orientation. In the attention conditions, strongest motion aftereffects were obtained when the orientations of the test stripes were orthogonal to the local motion directions which attention modulated, with both monocular and inter-ocular viewings. In the non-attention condition, the aftereffect was strongest when the orientation of the test stripes was orthogonal to the integrated motion direction (the rightward direction). These results show that motion aftereffects are modulated by attention to a local motion.

Second, in both the attention and non-attention conditions, the motion aftereffects were stronger with monocular viewing than with inter-ocular viewing. The difference in strength of the aftereffect between monocular and inter-ocular viewing in the non-attention condition may be attributed to the monocular units' activation, stimulated by the afferent stimulus information only with the monocular viewing (Wolfe & Held, 1981). However, the difference in the strength between the two viewings around the peak of the "tuning curves" was much larger in the attention conditions than in the non-attention condition. Furthermore, while the "tuning curve" for monocular viewing is as shallow as for inter-ocular viewing in the non-attention condition, the tuning curves were significantly steeper for monocular viewing than for inter-ocular viewing in the attention conditions.

These results suggest that the difference in strength and shape of the two curves in each attention condition may not be attributable just to the difference between the presence and absence of the adaptation of the monocular units to the afferent motion information. Presumably, the monocular units for attended motion directions were strongly activated. Attention seems to influence monocular units for relevant component motion directions.

Since a majority of monocular cells have been found in and before V1, these psychophysical results are in accord with the higher activation of V1 in the attention conditions with f-MRI. Furthermore, since the psychophysical results suggest the activation of a specific unit for relevant motion directions, they also eliminate the possibility that the activation of V1 found by f-MRI technique is due to the general arousal of V1.

The role of attention in the second stage of motion processing

As mentioned in the introduction, it has been suggested that locally measured motion signals interact and integrate at MT (Rodman & Albright, 1989; Movshon et al, 1986; Snowden, 1991).

In the previous section, we showed that attention to a local motion activates the units in V1 which correspond to local component computations (Movshon et al, 1986).

An important question arises: Does attention to any kind of motion always activate V1 or does the stage of motion processing that is activated by attention depend on the feature of motion to which attention is directed?

We examined the blood flow change by means of the f-MRI technique when attention was directed to an integrated direction of motion, instead of a local motion direction. In the experiment, the subjects were presented with several bouncing balls (Figure 6(a)).

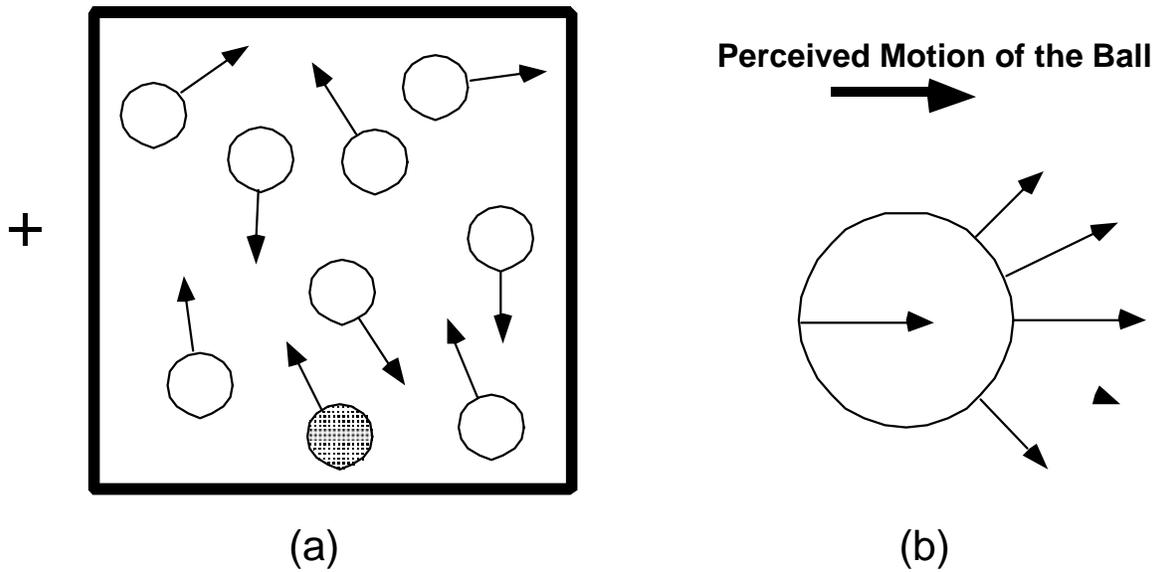


Figure 6: (a) Balls bouncing within a framework were presented. In the attention condition, the subject was instructed to direct attention to particular ball which blinked for a second before the onset of the measurement of the brain activity. In the passive condition, the subject was instructed to direct attention to no particular place. In both conditions, the subject had to fixate the fixation point which was presented either the right or the left side to the framework. (b) The perceived motion (represented by the bold arrow) of a bouncing ball is obtained as a result of integration of local motions whose directions are perpendicular to local contours.

In the attention condition, the subject was asked to direct attention to one particular ball while fixating a point in the center of the display. In the non-attention condition, the subject was asked to watch the display passively while fixating the central point. As shown in Figure 6(b), the perceived motion direction of a bouncing ball should be obtained by integrating locally measured motion signals. Thus, directing attention at a single bouncing ball among many balls means directing attention to an integrated motion among many integrated motions. We found that MT/MST area was significantly more activated in the attention condition than in the non-attention condition. However, there was no significant difference found in the amount of activation in V1 between the two conditions (Miyauchi et al, 1995, 1996). Basically the same results were obtained when attention was directed to expanding/contracting motion that may be first measured in MST (Tanaka, Fukada & Saito, 1989) as a result of interactions and integrations of local motion units (O'Craven and Savoy, 1995).

Further study

We are now conducting a series of experiments by means of f-MRI and psychophysics in which the subject is instructed to direct attention to a different feature of motion with the same stimulus under each of several different conditions. Although this experiment is still underway, the preliminary results show that the location of an activated area due to an attention allocation depends on a task demand, that is, which feature of motion attention is directed to.

Conclusion for effects of attention on motion processing

The results of the experiments have the following implications: First, when attention is directed to a component motion direction, V1 is more activated. Second, the V1 activation may be attributed to the activation of the units in V1 for the local motion, but not to an overall activation in V1 caused by a general awareness. Third, when attention is directed to an integrated motion of an object, MT is more activated, but V1 remains at baseline. That is, the level of motion processing attention influences depends on features of motion to which attention is directed.

Why does the stage of motion processing that is influenced by attention vary under different experimental conditions? Remember that according to the two stage model, the earliest stage at which local motions are measured is V1 while the area at which the local motion signals interact and are integrated may be MT/V5 (Movshon et al, 1986; Rodman & Albright, 1989; Snowden et al., 1991). Our findings suggest that attention influences a cortical area that is crucial for measuring or producing signals of a feature of motion (local component motion or global motion) to which attention is directed.

This makes several important implications regarding the relationship of attention to visual information processing in general. First, basic visual features such as motion signals in V1 can be modulated by attention. Several researchers have argued that visual search for basic features are in a relatively high-level stage of visual processing (Cavanagh, Arguin & Treisman, 1990; Bravo & Blake, 1990). Perhaps attention can reach down and modulate basic visual features in representations that come before the representation on which visual search occurs (Wolfe, 1996). Second, our results indicate that the units for a relevant motion feature are activated. This suggests that attention enhances signals for a relevant feature, rather than just picking up the signals and/or filtering out irrelevant signals. Third, the visual system seems to activate the specific unit in a specific stage which is most responsible for the processing of a feature at which attention is directed.

INTER-MODULAR INTERACTIONS BETWEEN MOTION PROCESSING AND FORM PROCESSING

The previous section indicated a way of sending feedback signals from a high-level cortical area influence early motion processing. In this section, we will show a novel phenomenon called *velocity decomposition* that may show how inter-cortical interactions occur— specifically, how motion processing interacts form and other kinds of processing to produce a unified surface representation or representations that are consistent with each other (Watanabe, 1997).

Watanabe (1997) showed that a transparency configuration makes two different interactions of motion signals occur in the same retinal region. He demonstrated that motion in one physical direction appeared to be decomposed into two component motions in a transparency configuration. Figure 7 (a) is a physical velocity map of the display that induces the velocity decomposition. It consists of random dots moving simultaneously in three different directions in three different regions—a central square and two flanking rotated "L" shapes. In this case, the velocity of the dots in the central square appeared to be decomposed into two component velocities; one identical to the velocity of the dots in the upper "L" and the other identical to the velocity of the dots in the lower "L", as illustrated in Figure 7 (b). In addition, one square with dots moving upward and another with the dots moving rightward appeared to overlap transparently in the central square⁵.

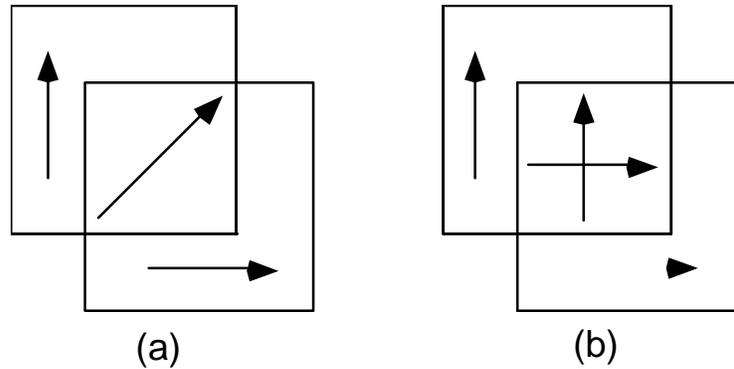


Figure 7: (a) Example of the stimulus that induces velocity decomposition. The three regions are defined by the different velocities of random dots. There are no black contour lines actually present as used in this figure. A velocity of the dots within the same area is represented as one vector. (b) Schematic description of perceived velocities. The velocity of the dots in the central square appears to be decomposed into two component velocities; one is the same as the velocity of the dots in the top “L” and the other is the same as that in the bottom “L”.

Watanabe (1997) checked how these dots from the three regions interact. In an experiment, a fixation point was presented for 1 sec. From trial to trial, its location was varied from the center of the central square to a point 4.5 deg above the center. Then the test stimulus (Figure 7a) was presented for one sec. While the dots in the upper and lower “L”s constantly moved rightward and upward, respectively, the direction of dots in the

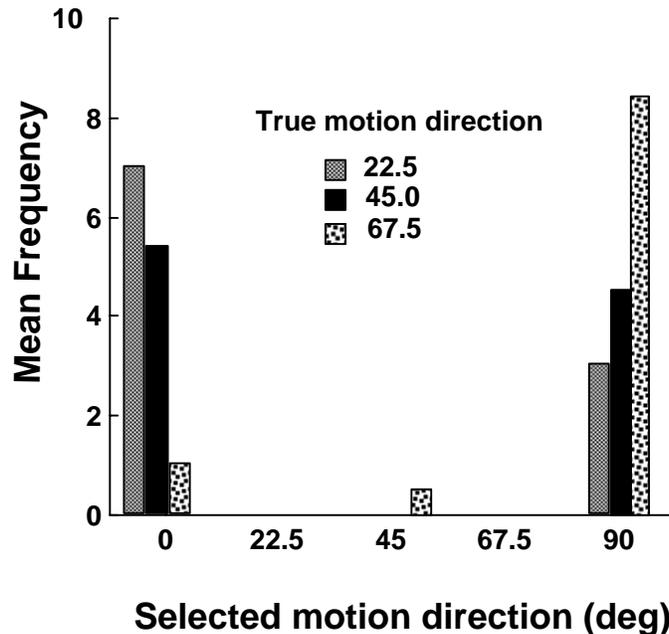


Figure 8: Mean frequency (n=2) as a function of the selected motion direction, for 22.5, 45.0 and 67.5 physically true motion directions of the dots in the central square. 0, 45, 90 deg represent the upward, upper-rightward, and rightward motion directions.

central square was varied from trial to trial in 16 different directions ranging from 0 to 337.5 deg in 22.5 deg steps. 0, 90.0, 180.0, and 270.0deg represent upward, rightward, downward and

leftward motion directions, respectively. The test stimulus was then followed by a set of arrows, each of which pointed in the 16 different directions. Individual dots existed for about 70 ms and then disappeared to be replaced by others at random locations so that the subjects could not trace individual dots to infer their physically true motion directions. This renewal scheme minimizes the probability that individual dot paths are being tracked (Andersen et al, 1990). The subjects were instructed to choose only one arrow that represented the perceived motion directions in the central square most appropriately.

In order to examine which motion direction the subjects saw for 22.5, 45.0 and 67.5 deg physical motion directions of the dots in the central square, the frequency of the motion direction the subject chose is shown for these three directions in Figure 8, respectively. The data were obtained when the location of the fixation point was 4.5 deg above the center of the central square. For 45 deg, they choose 0 and 90 deg nearly equally often. This result is in accord with the subject's verbal report that they actually saw 0 and 90 deg motions at the same time. However, since they were instructed to choose only one motion direction, they chose one of the two randomly. On the other hand, when the physical true motion directions were 22.5 and 67.5 deg, they saw 0 and 90 deg motion directions most frequently. What was common to all the three conditions was that with increasing eccentricity from the center of the central square, the frequency of the subject choosing 0 or 90 motion directions increased, although even with the foveal viewing the same but weaker tendency was observed.

Watanabe (1997) also conducted an experiment in which the speed of the dots in the central square of the same figure as in the previous experiment was systematically varied and found that velocity decomposition occurred most frequently when the velocity (direction and speed) of the dots in the central square is the vector sum of the velocities of the dots in the two "L"s. The same tendency was found with many other combinations of dot motion directions in the three areas.

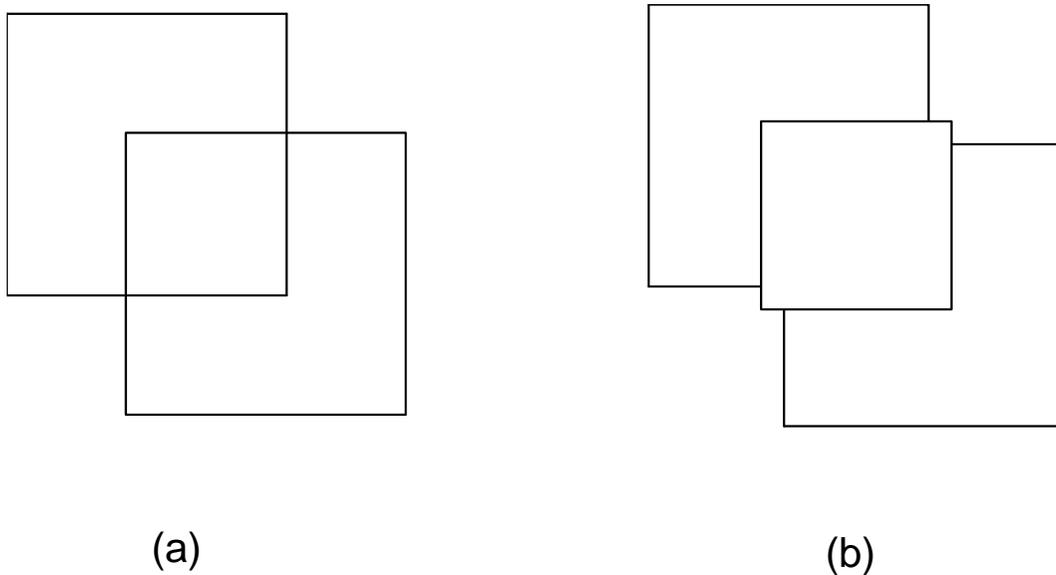


Figure 9: (a) A transparency configuration. The two large squares appear to be overlapping in the region of a central small square. (b) An occlusion configuration. The central square appears to be opaque, occluding the two large squares.

Velocity decomposition has been observed in the transparency configuration shown in Figure 9 (a), in which two large squares appear to be overlapping in the region of the central small square. At times, the top large square appears to be transparent, overlying the bottom square. At others,

the bottom appears to be transparent. Is the transparency configuration necessary for velocity decomposition?

In order to answer this question, the random dots were put in an occlusion configuration as shown in Figure 9(b), where the central square appears to be opaque, partially occluding the two large squares. In this case, no velocity decomposition was observed even if the velocity of the dots in the central square was the vector sum of the velocities of the two "L"s. The subject constantly chose 45 deg motion direction when this direction was presented in the central square. These results show that the transparency configuration is necessary for velocity decomposition³.

How do velocity and surface decompositions occur? As mentioned above, the motion condition necessary for velocity decomposition is also necessary for surface decomposition and the figural configuration necessary for surface decomposition is also necessary for velocity decomposition. Since surface decomposition occurs in the figural configuration shown in Figure 9a, form processing should be involved in the surface decomposition. Thus, the results of the present study suggest that there is a strong reciprocal interaction between motion and form processings in order to produce a unified representation presentation (Cavanagh, 1987) or presentations consistent with each other. The information of a representation in form processing may influence and change a representation of motion. This change in motion representation might, in turn, influence the representation in form processing, and so on. This reciprocal interaction may be accomplished by inter-module projections.

What is happening within motion processing in relation to the two stage motion model while motion processing and form processing interact? There are at least two possibilities. One is that an interaction occurs between form processing and the second stage of motion processing. Velocity decomposition may be a result of two different global interactions of motion direction signals occurring at two different depth planes in the same retinal region. For the 22.5 and 67.5 degs, the motion of the dots in the central square seems to be assimilated to the motion direction of an "L" which is closer to the direction of the dots in the central square (see, Nawrot & Sekuler, 1990; Watanabe & Cole, 1995). Assimilation may be the result of large-scale interactions such as smoothing of different motion directions (Bulthof, Little & Poggio, 1989 ; Grossberg & Rudd, 1992; Hildreth, 1984; Yuile & Grzywacz, 1988). When the velocity of the dots in the central square was the vector sum of the velocities of the dots in the "L"s, the magnitude of the difference between the motion directions of the dots in the upper "L" shapes and that in the central square was the same as the magnitude between those in the lower "L"s and in the central square. When this velocity condition and the transparency configuration coincide, the two different assimilations of the dots in the central square to the directions of the dots in the two "L"s could occur at two different depth planes into which the central square was decomposed. On the other hand, any assimilation did not occur in the occlusion configuration that does not induce surface decomposition.

The other possibility is that the local motion signals only for the decomposed motion directions are relatively enhanced in the first stage of motion processing. The direction of local motion is suggested to be orthogonal to the local contour (Gizzi et al, 1990; Worgotter, 1991). Usually, local motion signals are integrated into one motion direction if these signals are regarded as properties of a single object. The speed of each motion component is proportional to the cosine of the angle between the true motion direction of the object and the local motion direction (Adelson & Movshon, 1982). When a circular dot is moving, for example, in the upper-right direction, the speed for the detected local motion for the true motion direction is the highest and much higher than those for the upward and rightward directions. However, in the surface decomposition configuration, the speeds of the local motions for the upward and rightward directions detected in each of the dots in the central square might be enhanced and the local motion signals might be integrated into these two directions.

CONCLUSION

A research on motion processing has mainly attempted to clarify only feedforward processing within the motion module. Indeed, research inspired by this framework has made a great progress and two stage motion processing has been largely accepted. However, a large number of reciprocal projections have been found between different cortical areas.

This chapter has shown our attempts to understand motion processing from a viewpoint of visual processing as an interaction of interdependent processes. First, we have shown how attention works on each of the two stages in motion processing. Using f-MRI technique, we found that attention directed to a local motion direction enhanced activity in parts of V1, whereas attention activated MT but not V1 when it is directed to a global motion direction. These findings have important implications for the role of feedback projections. The visual system seems to control the destination of feedback processing so well as to activate a specific unit for motion direction up to a specific stage of motion processing.

Second, we showed a phenomenon called velocity decomposition that may reveal reciprocal interactions between motion and other processings whose implementation may be accomplished by inter-module projections.

Although only two examples are shown in this chapter, there may be many visual functions whose feedforward processing within a module are largely known but whose interactive aspects are a mystery. In the near future, more of these interactive aspects are expected to be clarified through research on visual processing from the viewpoint of visual processing as an interaction of interdependent processes.

FOOTNOTES

¹ Many computational models assume that the first stage can be divided into energy extraction through spatio-temporal filters and construction of a local velocity from the extracted energy (e.g., Adelson & Bergen, 1985; Grzywacz & Yuile, 1991; Heeger, 1987).

² There are two controversial hypotheses as to an integrated velocity. One is the "intersection-of-constraints" (IOC) hypothesis that assumes that an integrated motion is determined by the intersection of all the physically possible interpretations generated by locally measured component motions (Adelson & Movshon, 1982; Fennema & Thompson, 1979). The other is the vector summation hypothesis in which the integrated velocity is determined by summation of locally measured component motions that are orthogonal to the local contours (e.g., Mingolla, Todd & Norman, 1992; Yo & Wilson, 1992).

³ The same transparency configuration causes brightness transparency (Beck, 1986; Beck, Pradny, & Ivry, 1984; Watanabe & Cavanagh, 1993) and texture laciness (Watanabe & Cavanagh, 1996).

⁴ The way of how suppressive interactions occur was measured psychophysically, first, by Snowden (1990) and, more recently, by Verstraten, Fredericksen, van Wezel, Boulton and van de Grind (1996) with an improved measurement.

⁵ In contrast, Lu and Sperling (1995) suggest that attention influences a saliency map rather than motion measurement. The saliency map is assumed to be in a lower stage than motion measurement.

REFERENCES

Adelson, E. H., & Bergen, J. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2, 284-299.

- Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *30*, 523-525.
- Alais, D., Wenderoth, P., & Burke, D. (1994). The Contribution of One-Dimensional Motion Mechanisms to the Perceived Direction of Drifting Plaids and their Aftereffects. *Vision Research*, *34*, 1823-1834.
- Andersen, R. A., Snowden, R. J., Treue, S., & Graziano, M. (1990). Hierarchical processing of motion in the visual cortex of monkey. *Cold spring harbor symposia on quantitative biology*, *LV*, 741-747.
- Anstis, S. (1986). Motion perception in the frontal lobe. In K. R. Boff, L. Kaufman, & J. R. Thomas (Eds.), *Handbook of Perception and Human Performance, Vol. 1. Sensory Processes and perception* (pp. 1-27). New York: John Wiley & Sons.
- Beck, J. (1986). Perception of transparency in man and machine. In A. Rosenfeld (Eds.), *Human and machine vision II* (pp. 1-12). Orlando: Academic Press.
- Beck, J., Pradny, K., & Ivry, R. (1984). The perception of transparency with achromatic colors. *Perception Psychophysics*, *35*, 407-422.
- Bravo, M. & Blake, R. (1990). Preattentive vision and perceptual groups. *Perception*, *19*, 515-522.
- Bulthoff, H., Little, J. & Poggio, T. (1989). A Parallel algorithm for real-time computation of optical flow. *Nature*, *337*, 549-553.
- Cavanagh, P. (1991). Attention-based motion perception. *Science*, *257*, 1563-1565.
- Cavanagh, P., Arguin, M. & Treisman, A. (1990). Effect of surface medium on visual search for orientation and size features. *Journal of Experimental Psychology: Human Perception & Performance*, *16*, 479-192.
- Chaudhuri, A. (1990). Modulation of the motion aftereffect by selective attention. *Nature*, *344*, 60-62.
- Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1991). Selective and divided attention during visual discriminations of shape, color, and speed; functional anatomy by positron emission tomography. *Journal of Neuroscience*, *11*, 2383-2402.
- Culham, J. C. & Cavanagh, P. (1994). Motion capture of luminance stimuli by equiluminous color grating and by attentive tracking. *Vision Research*, *34*, 2701-2796.
- DoYoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neuroscience*, *11*, 219-226.
- Gizzi, M. S., Katz, E., Schumer, R. A., & Movshon, J. A. (1990). Selectivity for orientation and direction of motion of single neurons in cat striate and extrastriate visual cortex. *Journal of Neurophysiology*, *63*, 1529-1543.
- Gorea, A., & Lorenceau, J. (1991). Directional Performances with Moving Plaids: Component-related and plaid-related processing modes coexist. *Spatial Vision*, *5*, 231-252.
- Grossberg, S. (1997). How is a moving target continuously tracked behind occluding cover? In *High-level Motion Processing* (Ed. T. Watanabe), Cambridge: The MIT Press.
- Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading, *92*, 173-211.
- Grossberg, S. & Rudd, M. E. (1992). Cortical dynamics of visual motion perception: short-range and long-range apparent motion. *Psychological Review*, *99*, 78-121.
- Grzywacz, N. M., & Yuile, A. L. (1991). Theories for the visual perception of local velocity and coherent motion. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 231-252). Cambridge: The MIT Press.
- Heeger, D. (1987). A model for the extraction of image flow. *Journal of the Optical Society of America A*, *4*, 1455-1471.
- Hildreth, E. C. (1984). *The measurement of visual motion*. Cambridge: MA: MIT Press.
- Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology*, *148*, 574-591.

- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. , *160*, 106-154.
- Koch, C., & Davis, J. L. (Ed.). (1994). *Large-scale neuronal theories on the brain*. Cambridge: The MIT press.
- Lankheet, M. J. M., & Verstraten, F. A. J. (1995). Attentional modulation of adaptation to two-component transparent motion. *Vision Research*, *35*, 1401-1412.
- Lu, Z.-L., & Sperling, G. (1995). Attention-generated apparent motion. *Nature*, *377*, 237-239.
- Mingolla, E., Todd, J. T. & Norman, F. (1992). The perception of globally coherent motion. *Vision Research*, *32*, 1015-1031.
- Miyauchi, S., Watanabe, T., Sasaki, Y., Takino, R. & Putz, B. (1995). Voluntary attention to motion can specifically activate either V1 or MT. *The proceedings of 10th Tokyo Institute of Psychiatry International Symposium*, Tokyo.
- Miyauchi, S., Watanabe, T., Butz, B., Takino, R., & Sasaki, M. (1996). Voluntary attention to the motion of visually perceived objects can specifically active either V1 or MT. *Electroencephalography and Clinical Neurophysiology*, in press.
- Moran, J., & Desimone, R. (1985). Selective Attention Gates Visual Processing in the Extrastriate Cortex. *Science*, *229*, 782-784.
- Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2 and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, *70*, 909-919.
- Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1986). The analysis of moving visual patterns. In C. Chagas, R. Gattas, & C. Gross (Eds.), *Pattern recognition mechanisms* (pp. 117-151). New York: Springer Verlag.
- Nawrot, M., & Sekuler, R. (1990). Assimilation and contrast in motion perception: expolations in cooperativity. *Vision Research*, *30*, 1439-1451.
- O'Craven, K. M., & Savoy, R. L. (1995). Attentional modulation of activation in human MT shown with functional magnetic resonance imaging (fMRI). *Investigative Ophthalmology & Visual Science*, *36*, 3923.
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25-42.
- Rodman, H. R., & Albright, T. D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research*, *75*, 53-64.
- Sato, T. (1997). Dmax: relations to low- and high-level motion processes. In *High-level Motion Processing* (Ed. T. Watanabe), Cambridge: The MIT Press.
- Snowden, R. J. (1992). The perception of visual motion. *Current Opinion in Neurobiology*, *2*, 175-179.
- Snowden, G. R., Treue, S., Erickson, R. E., & Andersen, R. A. (1991). The response of Area MT and V1 neurons to transparent motion. *Journal of Neuroscience*, *11*, 2768-2785.
- Stoner, G. R., & Albright, T. D. (1992). neural correlates of perceptual motion coherence. *Nature*, *412*, 412-414.
- Stoner, G. R., Albright, T. D., & Ramachandran, V. S. (1990). Transparency and coherence in human motion perception. *Nature*, *344*, 153-155.
- Tanaka, K. (1997). Representation of visual motion in the extrastriate cortex. In *High-level Motion Processing* (Ed. T. Watanabe), Cambridge: The MIT Press.
- Tanaka, K., Fukada, Y., and Saito, H. (1989). Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, *62*, 642-656.

- Treisman, A., & Gelade, G. (1981). A feature integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Treue, S. & Maunsell, J. H. R. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*, 382, 539-541.
- Tse, P., Cavanagh, P. & Nakayama, K. (1997). The role of parsing in high-level motion processing. In *High-level Motion Processing* (Ed. T. Watanabe), Cambridge: The MIT Press.
- Yuille, A. L. and Grzywacz, N. M. (1997). A Theoretical Framework for Visual Motion. In *High-level Motion Processing* (Ed. T. Watanabe), Cambridge: The MIT Press.
- Verstraten, F. A. J., Fredericksen, R. E., & van de Grind, W. A. (1994). Movement aftereffect of bivectorial transparent motion. *Vision Research*, 34, 349-358.
- Verstraten, F. A. J., Fredericksen, R. E., van Wezel, R. J. A., Boulton, J. C. & van de Grind, W. A. (1996). Directional motion sensitivity under transparent motion conditions. *Vision Research*, in press.
- Watanabe, T. (1995). Motion aftereffects show that attention may selectively activate local motion units located in V1. *Investigative Ophthalmology & Visual Science*, 36, 634.
- Watanabe, T. (1997). Velocity decomposition and surface decomposition —reciprocal interactions between motion and form processing—, *Vision Research*, in press.
- Watanabe, T. & Cavanagh, P. (1993). Surface decomposition accompanying the perception of transparency. *Spatial Vision*, 7, 95-111.
- Watanabe, T. & Cavanagh, P. (1996). Texture laciness, *Perception*, in press.
- Watanabe, T., & Cole, R. (1995). Constraint propagation of apparent motion. *Vision Research*, 28, 2853-2861.
- Welch, L. (1989). The perception of moving plaids reveals two motion-processing stages. *Nature*, 337, 734-736.
- Wolfe, J. M., & Held, R. (1981). A purely binocular mechanism in human vision. *Vision Research*, 21, 1755-1759.
- Wolfe, J. M. (1996). Personal communication.
- Wurtz, R. H., Goldberg, M. E., & Robinson, D. L. (1980). Behavioral modulation of visual responses in the monkey: stimulus selection for attention and movement. In J. M. Sprague & A. N. Epstein (Eds.), *Progress in Psychobiology and Physiological Psychology*. (pp. 43-83). New York: Academic Press.
- Yo, C. & Wilson, H. R. (1992). Perceived direction of moving two-dimensional patterns depends on duration, contrast and eccentricity. *Vision Research*, 32, 135-147.
- Yuille, A. L., & Grzywacz, N. M. (1988). A computational theory for the perception coherent visual motion. *Nature*, 333, 71-74.
- Yuille, A. L. and Grzywacz, N. M. (1997). A Theoretical Framework for Visual Motion. In T. Watanabe (Ed), *High-level Motion Processing*, Cambridge: MIT Press.

環境に適応し，未来を予測する視知覚システム

西田眞也（NTT コミュニケーション科学基礎研究所）

1．はじめに

生体の感覚システムは定常的なものではなく，生体のおかれた文脈，つまり環境に応じて常に変化し続ける．そのことを自分の見えとして身をもって体験させてくれるのが順応現象である (Barlow, 1990)．順応とは，特定の刺激を観察し続けた（順応した）結果として，知覚が変容することをいう．順応現象は，知覚メカニズムを解析する有効な手段を我々に提供してくれる．さらに順応そのものが，知覚メカニズムの本質的な特性なのである．

2．運動順応現象

同じ方向に動き続ける刺激を見つめ続けると，その運動方向の刺激に対する検出感度が低下し，刺激検出できる刺激強度の下限，つまり検出閾が上昇する (Sekuler & Ganz, 1963)．また，同じ方向の運動を見続けたあとでは，止まっているものが反対方向にゆっくり動いているように見える (Mather, Verstraten, & Anstis, 1998; Wohlgenuth, 1911)．前者を運動方向選択的感度低下，後者を（静止）運動残効という．

視覚系には運動方向や速度に選択性のある細胞が数多く存在していて，運動センサーとしてはたらいている．特定の方向の運動を見せられると，それを最適刺激とする運動センサーが強く応答する．しかし，その状態が長く続くと「順応」のメカニズムによって次第に反応が弱まってくる．この神経レベルの感度低下が知覚に反映し，順応刺激に対する選択的な検出閾の上昇が生まれる．

運動残効も同じ原理で説明できる (Barlow & Hill, 1963)．一つの運動センサーの刺激選択性はそれほど鋭くないため，チューニングの異なる複数のセンサーの活動のパターンによって運動方向や速度は符号化されている．たとえば，静止刺激に対しては，どの運動方向のセンサーの反応も弱く，かつ同じ程度である．ところが，特定の方向の運動に順応し，強く応答したセンサーの感度が低下してしたあとでは，この関係は成り立たない．静止刺激に対して反対側の方向のセンサー群の反応が相対的に大きくなってしまふ．その結果，止まっているものが順応運動と反対方向に動いて見えるのである．

この二つの現象は順応後，検出閾レベルと閾上レベルの強度の刺激の知覚に現れる典型的な順応現象の姿を示している．一つは，刺激選択的な感度低下であり，もう一つは順応と反対方向への知覚のゆがみである．運動のほか，色，方位，空間周波数，両眼視差などの視覚次元や，聴覚などの他の感覚モダリティの知覚次元についても類似の現象がおこる．

順応現象は刺激選択性を持っている．すなわち，順応刺激とテスト刺激の類似性が高いほど，

順応効果が高い。順応時に感度変化したメカニズムがテスト刺激の知覚に関与しなければならないのだから、これは当然のことである。しかし、どういう次元で類似していなければならないかはアプリアリには決定できない。順応したメカニズムがどういう次元に選択性を持っているかによって、それは変わってくるからである。

逆に言えば、順応の選択性を調べれば、背後にあるメカニズムの特性が明らかになる。そのため、マスキングと並んで、順応は知覚システム解析のための代表的な心理物理学的手法とされてきた(Blakemore & Campbell, 1969; De Valois & De Valois, 1988)。そのことを運動視メカニズムについて具体的に説明する。

3. 運動順応現象による運動視システムの解析

運動方向選択的な閾値上昇は、刺激の網膜位置、方位、空間周波数にも選択性を示す。このことは、順応している運動センサーが、空間的に局在し、方位と空間周波数に選択性を持つことを示唆している。現在スタンダードとなっている運動エネルギーモデル(Adelson & Bergen, 1985)はまさにそのような受容野構造を持っている。また、近年の研究によって、モデルの生理学的妥当性は(大枠では)支持されている(De Valois & Cottaris, 1998; DeAngelis, Ohzawa, & Freeman, 1993)。

運動順応による閾値上昇は、入力眼に対する選択性も示す。右眼に順応刺激を提示し左眼にテスト刺激を提示すると、あまり閾値上昇は起こらない。つまり、両眼間転移が弱い。これは、順応している運動センサーが(一次視覚野の細胞でみられるように)部分的に単眼性であることを意味している。

さらに、運動方向選択的閾値上昇は刺激のタイプ(次数)について選択性がある。通常の運動では輝度(刺激の強度)の分布で成り立つパターンが移動する。これを一次運動という。これに対し、輝度分布の移動はなく、輝度より高次の刺激属性、たとえば2点の輝度関係で決まるコントラスト振幅で成り立つパターンだけが移動するものを二次運動という(Cavanagh & Mather, 1989; Chubb & Sperling, 1988)。この二次運動への順応は二次運動の検出閾を上昇させるが、一次運動の検出にはほとんど影響しない(Nishida, Ledgeway, & Edwards, 1997)。このような刺激のタイプ(次数)に対する選択性は、一次運動と二次運動が独立に検出されていることを示唆している。

運動エネルギーモデルは輝度の流れの抽出を想定したモデルであり、理論的に二次運動が検出できない。もし、一次運動と二次運動が共通のメカニズムで検出されているならば、運動エネルギーモデルは間違っていることになる。閾値上昇の結果は、この議論を支持せず、二次運動の検出に専用のセンサーがあることを示している。二次運動センサーは、二次の特徴を前処理で抽出してから運動を検出するメカニズムと考えられる。

次に、運動残効の特性である。閾値上昇と同じく、運動残効も刺激位置や空間周波数に選択性を示し、両眼間転移は部分的である。これは、閾値上昇に関わった運動センサーの順応(感度低

下)が運動残効の原因になっている可能性を示している。また、二次運動に順応した場合、運動残効は起こらない(Derrington & Badcock, 1985; Nishida & Sato, 1992)。一次運動への順応だけが運動残効を生み出すということも、一次と二次運動が別々に検出されているということの証拠となる。

しかし、これらはテスト刺激に静止刺激を用いる古典的な静止運動残効の特性である。止まった刺激が動いて見えるだけが運動残効ではない。運動方向が曖昧な運動刺激(たとえば位相反転格子縞)をテスト刺激に用いると、順応後しばらく順応と反対方向にだけ動いているように見える。これをフリッカー運動残効という。静止運動残効と同じく、フリッカー運動残効も運動センサーの順応で容易に説明できるが、その特性は静止残効と同じではない。

第1に、フリッカー運動残効は二次運動に順応したあとでも見える。さらに閾値上昇とも違って、二次運動に順応したとき一次のテスト刺激でも残効が見える(Ledgeway, 1994; Nishida & Sato, 1995)。第2に、フリッカー運動残効は空間周波数選択性が弱く、順応速度が重要な要因になる(静止残効の場合は、順応の時間周波数が重要)(Ashida & Osaka, 1994; Ashida & Osaka, 1995)。第3に、フリッカー運動残効の両眼間転移は高く、100%の転移を示すこともある(Nishida, Ashida, & Sato, 1994)。これらの特性は、フリッカー残効が一次と二次の運動情報を統合する高次システムの順応を反映していることを示唆している。

細かいことを言えば、フリッカー残効には高次と低次の成分が含まれていて、上記は高次成分についての特性である(Nishida & Ashida, 2000)。低次成分は、静止運動残効と同じように部分的な両眼間転移を示す。低次の成分は順応運動に注意を向けていないときに現れる。これは、高次の運動統合メカニズムが注意を向けないとはたらかないのに対して、低次の運動センサーは注意の修飾効果をあまり受けないからである。このように、トップダウンからの影響は階層性を持っている。

すべてをまとめると次のような運動視の構造が見えてくる。一次運動と二次運動は別個のセンサー群によって検出される。どちらのセンサー群も多重の空間スケールで運動検出し、(部分的に)単眼性の入力を受ける。二次運動の検出には非線形の前処理によって特徴を取り出すステージが含まれる。一次と二次の運動信号は上位で統合される。統合メカニズムは両眼性で、空間スケールの情報は統合され、速度の計算が行われる。高次統合メカニズムの処理対象は、視野の中心でとらえ注意を向けた運動刺激である。運動順応による運動方向選択的な閾値上昇は、低次の運動センサーの順応によって起こる。一次、二次の両方のセンサーが含まれる。静止運動残効は、低次の一次運動のセンサーの順応によって生じる。フリッカー運動残効は、低次の一次、二次の運動センサー、そして高次の運動統合メカニズムの順応効果を反映する。

4. 運動残効の形態知覚への影響

以上のように、運動順応現象を調べていくことで、運動視メカニズムの構造が明らかになる。さらに、運動残効を使えば、運動視メカニズムとその他の知覚メカニズムの関係も分析することができる。

視覚世界は、明るさや色、模様、形、運動、位置など、さまざまな画像の属性から構成されている。異なる属性を担当する機能は大脳皮質の別々の位置にある。例えば、運動の情報は五次視覚野から頭頂葉にいたる経路で処理されるが、色や形は四次視覚野から側頭葉に至る経路が深く関わっている。脳の中では別個の視覚属性がほぼ独立に処理されているのである。

では、それぞれの視覚属性の知覚は互いに影響しないで独立なのだろうか。例えば運動と位置、運動と形である。物理的な世界において、物体の運動は位置や形の変化を伴う。そして、一定時間でどれだけ位置が変化するか、から運動の速度が計算できる。運動の知覚とその際の位置や形の知覚はどのような関係にあるのだろうか。

静止運動残効においては、物理的に止まったものが動いて見える。このとき、テスト刺激の位置は変化して見えるのだろうか。それとも、同じ位置に見え続けるのだろうか。もし、運動と位置が全く別々に処理されているのならば、同じ位置に見え続けるはずだ。そして、まさにそれがこれまでの常識であった(Gregory, 1966)。しかし、その常識は間違っていたのである。

被験者にまず回転する風車図形を見せ、その後、テスト刺激となる静止風車を見せる。すると、運動残効によって被験者には風車図形が反対方向に回転しているように見える。ここで、本当は垂直である風車の矢羽根の方位が、被験者にどのように見えるかを調べる。運動残効が位置変化を伴わないなら垂直に見えるはずだが、実際には、運動して見える方向に傾いているように知覚されたのだ。さらに、この傾き錯視の時間的な変化を検討した結果、運動信号が時間的に積分されて位置信号に変換されていることが示唆された(Nishida & Johnston, 1999)。

この現象は、運動の情報が位置や形の知覚に影響するような経路が脳内に存在することを意味している。視覚系はそういう経路によって運動情報から次の瞬間に見えるべき情景の位置や形を予測することができる。このような時間的な補完機能によって、人間の脳は絶え間なく変化する複雑な視覚入力をリアルタイムで処理しているのである。

5. おわりに

順応現象が視覚系の解析の道具として非常に有効であることをみてきたが、最期に順応そのものがどうしておこるかについても少し触れておこう。

同じ刺激入力を受け続けたときそれに強く反応した神経センサーの反応が弱まることは、単なる疲労ではなく、適応的な意味を持っている。神経センサーのダイナミックレンジは狭く、入力信号のレンジに合わせてゲインをコントロールする必要があるからだ。つまり、高コントラストの刺激を見たあとで低コントラストの刺激が見えにくくなったのは、高コントラスト近傍での刺

激変化をとらえようとシステムがダイナミックレンジを調整して、低コントラスト刺激が対応レンジの外となったためである。その代わりに、高コントラスト付近のコントラスト変化に対してはむしろ感度がよくなる (Greenlee & Heitger, 1988) (ただし、この結果に否定的な報告もある)。同じように、運動刺激に順応すると、速度が遅く見えたり、止まっているものが反対に動いて見える一方で、順応した速度近傍での速度変化に対する弁別感度がよくなるのである (Clifford & Wenderoth, 1999)。

順応の機能は、現在の入力分布を最適に符号化するために、システムの特性を変化させることだと考えられている (Barlow, 1990; Wainwright, 1999)。現在の文脈によって内部モデルを更新し、過去から予想できることは折り込み済みとして、現状からの変化の検出に全力を傾けるとするのが知覚系の本質のようである。その本質が実験室的な状況においてあぶり出されたのが順応現象に他ならない。ただ、その計算論的な位置づけ (特に自己組織化や学習との関係) や神経インプレメンテーションの詳細は、今後検討すべき問題として残されている。

引用文献

- Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2, 284-299.
- Ashida, H., & Osaka, N. (1994). Difference of spatial frequency selectivity between static and flicker motion aftereffects. *Perception*, 23, 1313-1320.
- Ashida, H., & Osaka, N. (1995). Motion aftereffect with flickering test stimuli depends on adapting velocity. *Vision Research*, 35, 1825-1833.
- Barlow, H. B. (1990). A theory about the functional role and synaptic mechanism of visual after-effects. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 363-375). Cambridge: Cambridge University Press.
- Barlow, H. B., & Hill, R. M. (1963). Evidence for a physiological explanation of the Waterfall phenomenon and figural after-effects. *Nature*, 200, 1345-1347.
- Blakemore, C., & Campbell, F. W. (1969). On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology*, 203, 237-260.
- Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, 4, 103-129.
- Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A*, 5, 1986-2007.
- Clifford, C. W., & Wenderoth, P. (1999). Adaptation to temporal modulation can enhance differential speed sensitivity. *Vision Res*, 39, 4324-4332.

- De Valois, R. L., & Cottaris, N. P. (1998). Inputs to directionally selective simple cells in macaque striate cortex. *Proc Natl Acad Sci U S A*, *95*, 14488-14493.
- De Valois, R. L., & De Valois, K. K. (1988). *Spatial Vision*. New York: Oxford University Press.
- DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1993). Spatiotemporal Organization of simple-cell receptive fields in the cat's striate cortex. I. General Characteristics and postnatal development. *Journal of Neurophysiology*, *69*, 1091-1117.
- Derrington, A. M., & Badcock, D. R. (1985). Separate detectors for simple and complex grating patterns? *Vision Research*, *25*, 1869-1878.
- Greenlee, M. W., & Heitger, F. (1988). The functional role of contrast adaptation. *Vision Research*, *28*, 791-797.
- Gregory, R. L. (1966). *Eye and Brain*. New York: McGraw-Hill.
- Ledgeway, T. (1994). Adaptation to second-order motion results in a motion aftereffect for directionally-ambiguous test stimuli. *Vision Research*, *34*, 2879-2889.
- Mather, G., Verstraten, F. A. J., & Anstis, S. M. (1998). *The Motion Aftereffect: A Modern Perspective*. Cambridge, Mass: The MIT Press.
- Nishida, S., & Ashida, H. (2000). A hierarchical structure of motion system revealed by interocular transfer of flicker motion aftereffects. *Vision Research*, *40*, 265-278.
- Nishida, S., Ashida, H., & Sato, T. (1994). Complete interocular transfer of motion aftereffect with flickering test. *Vision Research*, *34*, 2707-2716.
- Nishida, S., Ledgeway, T., & Edwards, M. (1997). Dual multiple-scale processing for motion in the human visual system. *Vision Research*, *37*, 2685-2698.
- Nishida, S., & Sato, T. (1992). Positive motion after-effect induced by bandpass-filtered random-dot kinematograms. *Vision Research*, *32*, 1635-1646.
- Nishida, S., & Sato, T. (1995). Motion aftereffect with flickering test patterns reveals higher stages of motion processing. *Vision Research*, *35*, 477-490.
- Nishida, S. y., & Johnston, A. (1999). Influence of motion signals on the perceived position of spatial pattern. *Nature*, *397*, 610-612.
- Sekuler, R. W., & Ganz, L. (1963). Aftereffect of seen motion with a stabilized retinal image. *Science*, *139*, 419-420.
- Wainwright, M. J. (1999). Visual adaptation as optimal information transmission. *Vision Res*, *39*, 3960-3974.
- Wohlgemuth, A. (1911). On the aftereffect of seen movement. *British Journal of Psychology, Monograph Supplement*, *1*, 1-117.

聴覚系のダイナミクスと環境適応性

柏野牧夫 (NTT コミュニケーション科学基礎研究所)

1. はじめに

聴覚研究は、視覚研究ともまた異なる独特の発展をしてきた。聴覚末梢系における周波数分析や、音の大きさ・高さ・音色・空間定位などの要素的属性に関する研究は古くから盛んであった[1]。また、音声言語や音楽などに特化した処理過程の研究も相当の蓄積を持っている[2, 3]。ところが、生物の聴覚の最も基本的な機能、すなわち、周囲のどこで何が起きているかを音波から推定するという働きに関しては、ごく最近に至るまでほとんど注意を払われなかったと言ってよい。

しかし 1990 年代に入り、聴覚研究は急速に変貌を遂げてきた。まだ実環境での機能をのものを扱っているという段階ではないにせよ、実環境での機能を念頭に置いて、それを実現するための情報処理システムとして聴覚系を捉える視点が台頭してきた。その背景にはいくつかの要因を指摘できる。知覚現象の記述的研究においては、音の群化や分凝の問題が、複数音源が混在した音波を解釈する過程（聴覚情景分析[4]）という観点で捉え直され、さまざまな現象的法則が整理された。モデル的研究においては、視覚の計算論の影響も受けつつ、聴覚情景分析の知見を計算機に実装し、音源分離の問題などを解こうとする動きが盛んになった[5]。一方心理物理学では、刺激が広帯域かつ時間的に変化する複雑なものに変わってきた結果、聴覚系の非線形性や動特性がクローズアップされることとなった[1, 6, 7]。さらに神経科学では、研究の中心が末梢の蝸牛や聴神経から脳幹神経核、さらには聴覚野へと高次に移っていき、解剖学的な構造と生理学的な特性の解明が進むと同時に、脳活動の画像化による研究も進んだ[8, 9]。

かくして、刻々と変化する環境に即応して適切な行動をとるために、耳に到達する音波を分析して音源の定位と識別を行う情報処理過程としての聴覚系を、現象、機能、機構の各側面から総合的に解明するというのが、しばらくは聴覚研究の主流となるであろう。とりわけ問題となるのは、多くの場合解が一意に求まらないような音源定位および識別の問題に対していかに合理的な解を求めるか、雑音や伝送路による変形などの効果に対していかに頑健な処理を行うか、さらに、刻々と入力される大量の情報をいかに効率的に処理するかといった点である。

本稿では、音源定位に焦点を絞り、そのダイナミクスと環境適応性に関する最近の研究を紹介する。音源定位は、音源識別とならぶ聴覚の二大基本機能であるだけでなく、聴覚研究の中では、現象、機能、機構の各側面の解明が最もよく進んでいる分野でもある。

2. 音源定位の神経機構

本題に入る前に、音源定位に関わる聴覚系の解剖学的構造および生理学的な特性につい

て、左右方向の定位に絞って概説しておく[10]。

左右方向の音源定位の主要な手がかりは、両耳に到達する音響信号の時間差（interaural time difference; ITD）と音圧レベル差（interaural level difference; ILD）である[11]。聴覚系は、ITD と ILD を各周波数ごとに測定し、それらの情報を統合することによって音源の空間的位置を推定する（ただし ITD は主に低い周波数で、ILD は高い周波数で有効である）。

この計算は、並列階層構造をもつ神経経路によって実現される。まず、左右の蝸牛で周波数分析された信号は脳幹の上オリーブ核（哺乳類と鳥類では解剖学的名称が異なる場合があるが、ここでは原則として哺乳類での名称を採用する）で収斂し、周波数帯域ごとの ITD と ILD が、それぞれ別の部位で並列に検出される。このうち ITD は、Jeffress[12] が予言した遅延一致検出回路によって、また ILD は、左右耳からの興奮と抑制の割合によって、それぞれ検出される。

次の主要な処理段階は下丘の中心核である。ここでは、周波数別の ITD、ILD の表現が相互抑制などによってさらに先鋭化される。ここで処理経路が、脳幹の上丘に向かう経路と、内側膝状体経由で第一次聴覚野に向かう経路とに分岐する。後者はさらに連合野の各部位に連絡し、一部は前頭葉の前頭眼野に至る。

ここで重要な点がふたつある。第一に、上オリーブ核から下丘、第一次聴覚野と進む経路が基本的に周波数ごとの定位手がかりを検出・表現するものであるのに対し、上丘や前頭眼野などは、それらの情報が周波数間で統合された“空間そのもの”を表現する部位である。これらの部位は、眼球運動や体の動きの制御に関与しており、聴覚による空間情報もそのために用いられると考えられる。第二に、下丘上丘という脳幹レベルと、第一次聴覚野以降の皮質レベルとでは、情報の表現形態が異なる。前者では、特定の範囲の ITD、ILD、もしくは空間的位置に選択的に応答するニューロンが規則正しく配列したいわゆるマップ表現になっている。一方、後者では、似通った空間的応答特性をもつニューロン群がクラスタをなしており、隣接クラスタ間では応答特性にあまり相関はみられない。

3. 音源定位における文脈適応

従来の研究では、心理物理学的なものであれ生理学的なものであれ、単一の静止した音源を用いたものが大半を占めていた。しかし、日常の環境はもっと複雑かつダイナミックである。複数の音源が継時的あるいは同時に別の空間的位置で発音して複雑な刺激文脈を作り出したり、音源や聴取者が動いて両耳に入る音響信号が時間とともに変化したりすることも多い。そのような場合、音源定位は、単一静止音源の場合の単純な和ではなく、文脈依存的なものであることを示す証拠が集まりつつある。

筆者らは、音源定位における文脈適応について一連の心理物理学的研究を行い、いくつかの現象を見出した。第一に、ある ITD をもつ音の定位が、別の ITD をもつ音を聞いた直後には、反発方向にずれる（定位残効）[13]。この現象は、ITD の処理系が、絶対的な定位

よりも相対的な変化分の符号化を優先するように働くことを示している。第二に、異なる ITD をもつ 2 個の音を同時に提示すると、両者の定位が反発する方向にずれる（定位同時対比）[14]。この現象は、定位残効も含む対比効果が、きわめて速く動作する神経過程に起因することを示している。さらに、ITD の弁別閾は、先行する音の ITD の近傍では低下し、離れた部分では上昇する（定位弁別残効）[15]。したがって、顕著な音がある場合、その周辺に自動的に焦点が絞られ、空間の解像度が向上することになる。

これらの 3 種類の現象はいずれも周波数選択性と ITD 選択性をもつので、周波数別に ITD がマップ表現されている部位が関与していると考えられる。また、定位弁別残効の実験において、先行音の ITD は左耳が進んでいるが ILD は右耳が強いというように相反する ITD と ILD を先行音に与えて両者の効果を相殺すると、ITD の弁別閾は先行音の音像が知覚される中央ではなく、先行音の ITD の位置で低下した[16]。このことは、この現象が、ITD と ILD が統合されて空間の表現が形成された後の段階で生じるものではなく、ITD 処理機構の段階で知覚意識に関係なく自動的に生じるものであることを示唆している。

そこで筆者らは、周波数と ITD に選択性をもつニューロン群の感度（利得）が先行刺激に応じて変化するモデルをつくり、シミュレーションを行ったところ、上記の現象の多くの側面を説明することができた。これとは独立に、Cai らは、相互抑制と順応を取り込んだ下丘のモデルを提案し、さまざまな文脈効果を説明することを試みている[17]。哺乳動物の下丘を対象とした電気生理学的実験によって、これらのモデルを裏付けるデータが得られつつある[18-20]。

このように動的な利得制御が存在すると、時間的にどういう刺激が先行したか、あるいは空間的にどういう刺激が共存しているかによって、検出器の同調特性が変化することになる。このような適応符号化は、固定的符号化に比べて、情報処理上の利点を持つ。単一の神経素子のダイナミックレンジは限られているので、仮に全刺激範囲を固定的符号化で処理しようとするれば、分解能は低くならざるを得ない。適応符号化は、最近経験した刺激分布を考慮して、その範囲の刺激を最大の分解能で符号化できるように神経素子の再校正を行うことに相当する。これによって、実効的な分解能が上がり、耐雑音性も向上する。さらに、時間軸上・ITD 軸上の変化分が強調され、原信号に含まれる冗長性が低減されるので、情報処理の効率が上がる。

これまでに研究されているのは刺激に依存したボトムアップ的、自動的な文脈適応であるが、同様の利得制御がトップダウンの信号によって働けば、能動的な空間的注意の一端を説明することができるであろう。この可能性の検証は今後の興味深いテーマである。

4. 訓練と学習

前章で述べたのは数秒程度の短時間内の刺激文脈に対する音源定位システムの適応であったが、もっと長い時間スケールの中でも適応を考えることができる。例えば、ITD や ILD、さらには音源方向依存的なスペクトル変化などの音源定位の手がかりは、聴取者が成長し

て頭部の形や大きさが変わるとともに変化するが、聴覚系はそのような変化に適応して適切な定位判断を保つ。さらに最近では、成長後の数日から数ヶ月程度の訓練や学習によっても、音源定位の特性が変化することを示す実験結果が得られつつある。

筆者らは、ITD 弁別課題における訓練の効果を調べた[21]。その結果、特定の ITD について ITD 弁別の訓練を数日間行くと、その ITD では弁別閾が有意に低下することが示された。しかし、訓練しなかった ITD では、弁別閾は変化しなかった。この結果は、両耳処理系のうちでも、特定範囲の ITD の処理を司るニューロン群が学習に関与していることを示唆している。

一方、視覚における上下反転眼鏡への適応のように、音響信号と音源定位との対応関係を人工的に変化させた場合の適応についての興味深い実験結果も報告されている[22]。耳介のくぼみの形状を変化させた状態では、音源方向依存的なフィルタリングの効果が通常と変わるので、とくに上下方向の音源定位がきわめて不正確になる。しかし、約 1 ヶ月その状態で生活すると、ほぼ正しく定位できるようになる。これは、体性感覚や視覚からのフィードバックによって、脳内の空間表現が再校正されたためと考えられる。

このような経験による可塑性のメカニズムに関する神経科学的な知見が、メンフクロウで得られている[23-26]。メンフクロウの視蓋（上丘に相当）には、聴覚と視覚の空間マップが対応して配列されている。幼鳥が視野をずらすようなプリズム眼鏡をつけられて成長すると、視蓋の聴覚マップがそれに対応してシフトする。このようなシフトは、視蓋の中のみならず、その前段階の、聴覚の空間マップが最初に形成される下丘外側核でも生じている。また、幼鳥のときにプリズム眼鏡に対応した聴覚空間マップを獲得した場合、正常なマップも残っているが、それは抑制によって不活性化されている。

5. おわりに

以上のように、音源定位の神経機構は、短期的には刺激文脈に適応し、より長期的には経験によって変容するダイナミックなものであることがわかってきた。いずれも、刺激の確率分布を勘案し、それを最も効率よく符号化できるように神経系の特性を再校正することであり、それによって聴覚系の環境適応性の一端が実現されている。今後は、短期的な適応におけるトップダウン制御の役割、知覚学習におけるパフォーマンス向上のメカニズム、視覚や体性感覚、自己受容感覚なども含めたマルチモーダルな相互作用などの研究を中心に、聴覚系のダイナミクスの解明がさらに進展することが期待される。

文献

- [1] Moore, B. C. J. (1997). *An Introduction to the Psychology of Hearing*. (Fourth ed.). San Diego: Academic Press.

- [2] 柏野 牧夫 (2000). 音声知覚の恒常性. 脳科学ハンドブック, 朝倉書店.
- [3] Deutch, D. (1982). *The psychology of music*. San Diego: Academic Press.
- [4] Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge: MIT Press.
- [5] Ellis, D. & Cooke, M. (in press). Auditory scene analysis by humans and machines. *Speech Communication*.
- [6] 柏野 牧夫 (1998). 聴覚環境に適応する無意識の知性. 日本音響学会誌, 54, 508-514.
- [7] 柏野 牧夫 (2000). 聴覚の心理学的モデル. 脳科学ハンドブック, 朝倉書店.
- [8] Popper, A. N. & Fay, R. R. (Eds.) (1992). *The mammalian auditory pathway: Neurophysiology*. New York: Springer-Verlag.
- [9] Rauschecker, J. P. (1998). Cortical processing of complex sounds. *Current Opinion in Neurobiology*, 8, 516-521.
- [10] Cohen, Y. E., & Knudsen, E. I. (1999). Maps versus clusters: different representations of auditory space in the midbrain and forebrain. *Trends in Neuroscience*, 22, 128-135.
- [11] Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization*. (Revised ed.). London: MIT Press.
- [12] Jeffress, L. A. (1948). A place theory of sound localization. *Journal of Comparative Physiological Psychology*, 41, 35-39.
- [13] Kashino, M., & Nishida, S. (1998). Adaptation in the processing of interaural time differences revealed by auditory localization after effects. *Journal of the Acoustical Society of America*, 103(6), 3597-3604.
- [14] Kashino, M. (1999). Interaction in the perceived lateralization of two sounds having different interaural time differences. *Journal of the Acoustical Society of America*, 105, 1343.
- [15] Kashino, M. (1998). Adaptation in sound localization revealed by auditory after-effects. In A. R. Palmer, A. Rees, A. Q. Summerfield, & R. Meddis (Eds.), *Psychophysical and physiological advances in hearing* (pp. 322-328). London, UK.: Whurr Publishers.
- [16] Kashino, M. (2000). Effects of a preceding sound on interaural-time-difference (ITD)

discrimination: ITD or perceptual space? *Association for Research in Otolaryngology*.

- [17] Cai, H., Carney, L. H., & Colburn, H. S. (1998). A model for binaural response properties of inferior colliculus neurons. II. A model with interaural time difference-sensitive excitatory and inhibitory inputs and an adaptation mechanism. *Journal of the Acoustical Society of America*, 103, 494-506.
- [18] Spitzer, M. W., & Semple, M. N. (1991). Interaural phase coding in auditory midbrain: Influence of dynamic stimulus features. *Science*, pp. 721-724.
- [19] Spitzer, M. W., & Semple, M. N. (1993). Responses of inferior colliculus neurons to time-varying interaural phase disparity : Effects of shifting the locus of virtual motion. *Journal of Neurophysiology*, 69(4), 1245-1263.
- [20] McAlpine, D., Jiang, D., Shackleton, T. M., Palmer, A. R. (2000). Responses of neurons in the inferior colliculus to dynamic interaural phase cues: evidence for a mechanism of binaural adaptation. *Journal of Neurophysiology*, 2000, 83, 1356-65.
- [21] Kawashima, T., Kashino, M., & Sato, T. (2000). Perceptual learning in the discrimination of interaural time differences. *Journal of the Acoustical Society of Japan (E)*.
- [22] Hofman, P. M., van Riswick, J. G. A., & van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nature Neuroscience*, pp. 417-421.
- [23] Brainard, M. S., & Knudsen, E. I. (1993). Experience-dependent plasticity in the inferior colliculus: A site for visual calibration of the neural representation of auditory space in the barn owl. *The Journal of Neuroscience*, 13, 4589-4608.
- [24] Brainard, M. S., & Knudsen, E. I. (1995). Dynamics of visually guided auditory plasticity in the optic tectum of the barn owl. *Journal of Neurophysiology*, 73, 595-614.
- [25] Zheng, W., & Knudsen, E. I. (1999). Functional selection of adaptive auditory space map by GABA mediated inhibition. *Science*, 284, 962-965.
- [26] Gold, J. I., & Knudsen, E. I. (2000). A site of auditory experience - dependent plasticity in the neural representation of auditory space in the Barn owl's inferior colliculus. *The Journal of Neuroscience*, 20, 3469 - 3486.

5 日目 階層的な予測 高次運動野と前頭連合野

コーディネータ：池田思朗，春野雅彦

1. ねらい

運動野には高次から低次への結合だけでなく双方向性結合や大脳基底核，小脳を含むループ結合が存在することが解剖学的に知られている．一方，近年の電気生理学的研究によりそれぞれの運動関連領域における情報処理と表現の変換に関する知見が蓄積されてきている．ここでは双方向情報処理，隠れ変数推定，モジュラーネットワークといった理論的手法と電気生理学的知見を組み合わせることにより運動関連領域での情報処理をシステムレベルで考察する手がかりを提供したい．

2. 講義内容/関連

午前前半の講義

階層的な情報処理システムにおいては入出力層以外の変数は外部から直接観測出来ないことが多い．この変数を隠れ変数と呼び，隠れ変数の導入によって一見異なる多くの問題を統一的に解くことが可能となる．この講義では階層的学習への情報論的理解を深めることを目的に，まず隠れ変数の考え方を導入し，続いてその推定法，双方向結合を持つニューラルネットワーク，他の統計モデルを用いた具体例などについて述べる．

午前後半の講義

この講義では目標軌道から制御コマンドを計算する逆モデルと制御コマンドからその結果を予測する順モデルをペアとしたモジュールを用いた階層的な運動学習モデルについて説明する．各モジュールは順モデルの予測の良さに応じて分化することで次第に要素運動と対応する．更に，高次のモジュールでは特定のシーケンスに特有の発火パターンを持つなど運動制御のトルクに密着した低次のモジュールとは異なった情報表現を獲得する．講義ではこれらのモジュールと脳との可能な対応や，モデルから考えられる実験についても述べる．またこのモデルの学習は 5 日目午前前半に説明がある隠れ変数推定の一種となっているので理論の適用例を知ることにもなる．

午後の講義

午後の講義では運動制御における座標の階層性について論じる．具体的には手首の単関節運動について空間，関節，筋肉の 3 つの座標系を分離する実験手法を開発し，前運動野と一次運動野における座標系を分析した．その結果，前運動野における空間座標系での運動指令が，主に一次運動野で筋肉座標系へと変換される階層的な機構が示唆された．講義ではこの機構を中心に，頭頂連合野，脊髄を含むより大きなシステムへの拡張を試みる．さらにこれらの皮質領域に見られる強力な双方向結合の意義を座標系の階層性という立場から考える．

3. 必読文献

Shiro Ikeda, Shun-ichi Amari, and Hiroyuki Nakahara, "Convergence of The Wake-Sleep Algorithm", In Michael S. Kearns, Sara A. Solla, and David A. Cohn editors, *Advances in Neural Information Processing Systems 11*, pp.239-245. The MIT Press, 1999

Masahiko Haruno, Daniel Wolpert, and Mitsuo Kawato, "Multiple Paired Forward-Inverse Models for Human Motor Learning and Control", In Michael S. Kearns, Sara A. Solla, and David A. Cohn editors, *Advances in Neural Information Processing Systems 11*, pp.31-37. The MIT Press, 1999

Shinji Kakei, Donna S. Hoffman and Peter L. Strick, "Muscle and Movement Representations in the Primary Motor Cortex", *Science*, vol.285, pp.2136-2139, 1999

隠れ変数と階層的学習

池田 思朗

科学技術振興事業団 さきがけ研究 21

埼玉県和光市広沢 2-1 理化学研究所 脳科学総合研究センター

Shiro.Ikeda@brain.riken.go.jp

1 はじめに

ここでは、確率分布における隠れ変数の考え方を説明する。隠れ変数とは、外からは直接観測できない確率変数である。

一つ例を考えよう。0, 1, 2, ..., 9 の数字が書かれた紙が沢山あるとしよう。そのうち一枚をとってきたとする。我々はこれを見て、今書かれている数字が 1 なのか 7 なのか、または 4 なのか、10 のカテゴリーのうちの一つに分類したい。

どの数字が表現されているのかを確率変数として考えてみる。すると、これは直接は観測できない確率変数であることが分る。我々は数字の書かれた紙を通じて何が書かれているのか推定しなければいけない。



図 1: 活字の場合



図 2: あいまいな数字

数字を見る前は、それぞれが等確率で出てくると予測しても構わないだろう。仮りに図 1 のようなデータを観測したのなら、ほとんど迷うことなく数字きた数字を当てることができる。

では、図 2 のような図が現れた場合はどうだろう。これはおそらく 1 だろうと考えられるが、7 である可能性も拭えない。つまり、数字を書いた人がどう思って書いたのかは直接は観測できないのである。ここまで書くと、この問題に隠れ変数を用いる意味が分るだろう。

実際の文字認識装置などを構成する場合、上のような問題をどうやって解決するのかは難しい問題であり、ここでは論じない。その代わりに、このような問題を情報量、確率を使って見るとどうなるのかを説明する。結果として、見通しの良い議論ができるようになると思う。

2 エントロピーと情報量

2.1 エントロピー

ある確率変数 X を考える。この確率変数の密度関数が $p(X)$ で与えられるとする。このとき、エントロピーは、

$$H(X) = - \int p(x) \log p(x) dx \quad (1)$$

と定義される。 X が離散値を取る場合には積分記号は単に和となり、

$$H(X) = -\sum_x p(x)\log p(x) \quad (2)$$

と定義される。

数字の例で考えると、10個の数字が等確率で出現するならば、 X をどの数字が出現するかという確率変数と置くと、

$$\begin{aligned} p(x=0) &= \frac{1}{10}, \dots, p(x=9) = \frac{1}{10} \\ H(X) &= -\sum \frac{1}{10} \log \frac{1}{10} = \log 10 \end{aligned}$$

となる。エントロピーとはある確率変数のもつ「曖昧さ」と考えてみても良い。この場合、どの数字が書かれているかという確率変数 X の持つエントロピー、すなわち曖昧さは $\log 10$ である。

2.2 相互情報量

エントロピーの考え方について、もう少し説明を続ける。

今、図1を観測したとする。ここで、新たな確率変数 Y を考える。 Y はどのような絵が観測されたかを示す確率変数である。例えば図1を観測したのであれば、 Y は図1となる。 Y は様々な図を取り得るので、確率変数として考えても良いだろう。

Y を観測したときの X の分布はどうなるだろうか。これは X の Y に対する条件付き確率 $p(X|Y)$ である。例えば図1の場合には書き手の意図としての数字は、確率1で「3」であろう。

これを確率で書くならば、

$$\begin{aligned} p(x \in \{0, 1, 2, 4, 5, 6, 7, 8, 9\} | y = \boxed{3}) &= 0, \\ p(x = 3 | y = \boxed{3}) &= 1. \end{aligned} \quad (3)$$

すると、この図を観測したときのエントロピーは、

$$H(X|y = \boxed{3}) = 0 \quad (4)$$

となる(計算中、 $0\log 0 = 0$ とした) すなわち、図を観測することで、曖昧さが消えエントロピーが0となり、結果として、 $\log 10$ であったエントロピーが0となった。このとき、 $\log 10 - 0$ の“情報量”が得られたことになる。

一方図2のような曖昧な図を観測したときは、

$$H(X|y = \boxed{7})$$

は0とならないだろう。どのように判断するかは別にして、これが1である確率が90%で、7である確率は10%とするならば、

$$H(X|y = \boxed{7}) = -\frac{1}{10} \log \frac{1}{10} - \frac{9}{10} \log \frac{9}{10} = \log 10 - \frac{9}{10} \log 9 \quad (5)$$

となる。得られた情報量は $\log 10$ との差であるから、 $\frac{9}{10} \log 9$ となり、依然として残っている曖昧さは $\log 10 - \frac{9}{10} \log 9$ である。

このように観測される図 Y と隠れた確率変数 X の間には相互に関係している。このような2つの確率変数の間に、相互情報量を考えることができる。 Y の密度関数を $q(Y)$ とすると、相互情報

量は Y を観測したときに、平均としてどのくらい情報量を得られるかとして定義できる。

$$\begin{aligned} I(X, Y) &= H(X) - \int q(y)H(X|y)dy \\ &= H(X) - H_Y(X) \end{aligned} \quad (6)$$

ここで、

$$H_Y(X) = \int q(y) \left[-\int p(x|y)\log p(x|y)dx \right] dy \quad (7)$$

と定義した。 $p(X)$ と $q(Y)$ が定義され、さらに $p(X|Y)$ が定義されれば X と Y の同時分布の密度関数を $r(X, Y) = p(X|Y)q(Y)$ として定義できる。これを用いると、

$$\begin{aligned} I(X, Y) &= H(X) - H_Y(X) \\ &= H(X) + H(Y) - H(X, Y) \\ &= H(Y) - H_X(Y) \end{aligned} \quad (8)$$

と書ける。ここで、

$$H(X, Y) = -\int r(x, y)\log r(x, y)dx dy \quad (9)$$

とした。この結果から相互情報量は X, Y に関して対称であることがわかる。

エントロピーや相互情報量を短い紙面で考えてみたが、ここに書いてあることのみでは十分ではない。情報理論の教科書を参考にして欲しい [1, 2]。

3 最尤推定

今までは、確率分布が定義されている上で、いくつかの事柄について説明をした。ここでは、ある程度確率分布が分っているが、そのパラメータが未知の場合、パラメータをどうやって求めるかについて、特に統計的な推定法である最尤推定を説明する。

例えば、データが正規分布に従っているということは既知だとする。しかし、平均値と分散が幾つかは分からないとする。このときに、データからパラメータをどうやって推定するかといったことにあたる。

θ をパラメータとする確率分布 $p(y; \theta)$ を考える。データが同じ確率分布から、独立に T 個 $\{y_1, y_2, \dots, y_T\}$ として得られたとする。このとき、 θ を推定したい。最尤推定では、 $p(y; \theta)$ がそのデータを受けとる確率(尤度)を最大にするパラメータを推定量 θ^* とする。

$$\theta^* = \operatorname{argmax}_{\theta} \prod_s p(y_s; \theta) = \operatorname{argmax}_{\theta} \sum_s \log p(y_s; \theta) \quad (10)$$

正規分布であれば、 $\theta = (\mu, \sigma^2)$ (平均と分散)であり、

$$p(y; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{\sigma^2}}$$

である。対数尤度は

$$\sum_s \log p(y; \theta) = -\sum_s \frac{(y_s - \mu)^2}{\sigma^2} - \frac{T}{2} \log 2\pi\sigma^2$$

となり、これを最大にする μ, σ^2 は

$$\mu^* = \frac{1}{T} \sum_s y_s, \sigma^{2*} = \frac{1}{T} \sum_s (y_s - \mu^*)^2$$

となる。

ここで、便利のために経験分布を $q(y)$ と定義する。例えば y が離散値を取るのであれば、

$$q(y) = \frac{1}{T} \sum_{i=1}^T \delta_{y_i}(y)$$

とすれば良い。 $\delta_{y_i}(y)$ は $y = y_i$ のときに 1 を取る関数である。

2つの確率分布の間の Kullback-Leibler Divergence は(11) のように定義される。 Kullback-Leibler Divergence はお互いの確率分布が一致したときのみ 0 になり、それ以外は正の値を取る。したがって分布間の違いを表わす。最尤推定は経験分布 $q(y)$ と $p(y; \theta)$ との間の Kullback-Leibler Divergence を最小にするようにパラメータを求めるのだと考えられる。

$$\begin{aligned} D(q, p(\theta)) &= \int q(y) \log \frac{q(y)}{p(y; \theta)} dy \\ &= \int q(y) \log q(y) dy - \int q(y) \log p(y; \theta) dy \end{aligned} \quad (11)$$

右側の式の第 2 項は(10) 式の対数尤度と等しい。パラメータに関する部分はこの項だけなので、対数尤度を最大にすることは(11) 式の量を最小にしていることと同値となる。

この結果を情報幾何 [3] を用いて解釈する。図 3 はこのイメージを示したものである。図中の S は y の確率分布の空間を考えたものである。この空間中の各点は y の確率分布となる。モデルは θ というパラメータを持つ集合であるので、この空間中では多様体 M として表わされている。経験分布を得たとき、ここからパラメータを最尤推定するとは、経験分布の 1 点 $q(y)$ からモデル多様体 M への一種の射影だとみなせる。この場合の射影は $D(q(y), p(y; \theta))$ を最小とする点を求めることと等しい [3]。

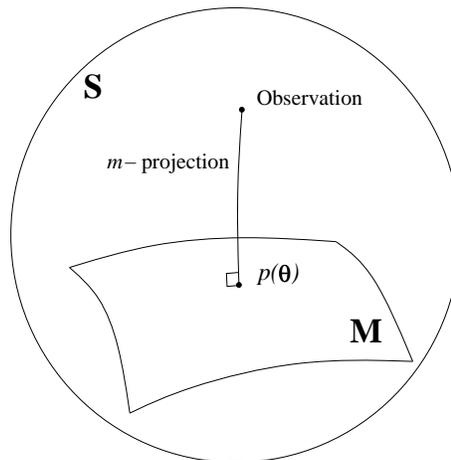


図 3: 統計的推定の幾何学的イメージ

4 隠れ変数を持つモデルとパラメータ推定

4.1 隠れ変数を持つモデル

はじめに隠れ変数について触れたが、ここでは幾つかの具体例を考える。まず、数式での扱い方を示し、例を交えて説明する。また、このようなモデルの最尤推定についても説明する。

Y として、観測できる確率変数があったとする。一方、変数 Z は観測できないとする。 Y, Z を両方含むモデルのパラメータを θ と書けば、全ての確率変数を含む確率密度関数は $p(y, z; \theta)$ と表わせる。しかし、本当に観測できるのは観測データ (y_1, y_1, \dots, y_n) のみであり、我々が観測データから得られる密度関数は、

$$p(y; \theta) = \int p(y, z; \theta) dz$$

という確率変数 Y についての周辺分布に関するものだけである。

例えば、例えば、上のような $\{0, 1\}$ からなる図があり、Fig.4 の5つの図のどれかが確率 $1/5$ で

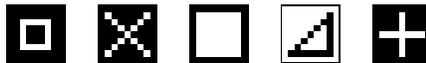


図 4: True images

取られるものとする。しかし、実際に提示される場合にはノイズが乗り、ビットがある確率 p_z で反転する。例えば、のような図が与えられるとする。我々が目にするのはこのようにノイズの



図 5: Sample images

乗った図のみであるとする。

この問題では、

隠れ変数 z : 実際に選択されている図は    のうちのどれかという確率変数。

観測される変数 y : 実際に観測されるノイズに汚れた図  のような図。

となる。モデルでのパラメータは $\theta = \{ \text{square with X}, \text{square with diagonal}, \text{square with plus}, p_z \}$ である。このパラメータを用いて尤度がどのように定義されるかを見よう。図5のような図が得られたのならば、

$$p(y; \theta) = \sum_z p(y, z; \theta) = \sum_z p(z; \theta) p(y|z; \theta)$$

から、

$$p(\text{noisy square with X}; \theta) = \frac{1}{5} p(\text{noisy square with X}; \theta) + \frac{1}{5} p(\text{noisy square with X}; \theta) + \dots + \frac{1}{5} p(\text{noisy square with X}; \theta)$$

となる。

このように例え $p(y; \theta)$ を知ったとしても、 Y のみからでは z を直接は知ることができない。そこで Z は隠れ変数 (Latent Variable) と呼ばれる。統計学で用いられる確率分布のなかには、混合分布や因子分析のモデルのように隠れ変数を持つものが多くある。また、脳のモデルとして提案されているニューラルネットワークにおいても隠れ変数の考え方を用いるものがある。次節以降でいくつかの例を示す。

4.2 混合正規分布

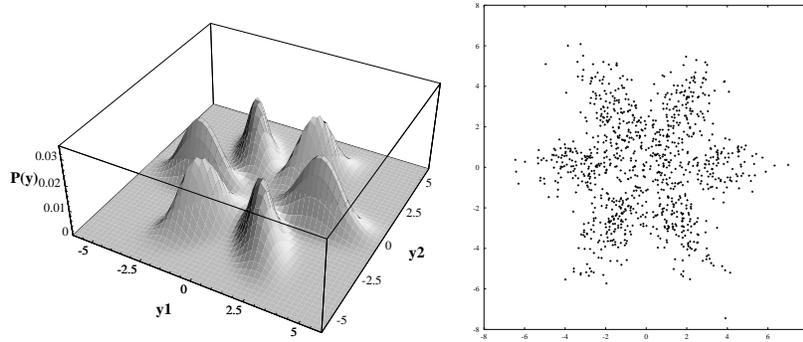


図 6: 正規混合分布

まずは、混合正規分布を考えよう。確率変数 Y の平均が μ で共分散行列が Σ の多次元正規分布にしたがうとき、その密度関数を $G(\mathbf{y}; \mu, \Sigma)$ と書くことにする。混合正規分布はこの正規分布の混合分布として表わされる分布である。密度関数 $p(\mathbf{y}; \theta)$ (ただし θ はモデルのパラメータ) は、ある π_i ($\sum_i \pi_i = 1$) を重み係数として、

$$p(\mathbf{y}; \theta) = \sum_i \pi_i G(\mathbf{y}; \mu_i, \Sigma_i)$$

となる。このモデルは隠れ変数を持つ。ではその隠れ変数はなんだろうか、

2次元の正規分布が6つ重なった正規混合分布を例に考えよう。確率分布の形を図6左に示す。この確率分布からデータが得られているとする。データを2次元平面に表示したものを図6右に示す。この場合、出力されるデータのみからでは、そのデータがどの正規分布によるものかははっきりとは分らない。すなわち、どの正規分布からのデータかという情報は観測できない。この「どの正規分布から発生したか」という情報が隠れ変数となる。正規混合分布の場合は隠れている確率変数は離散的な確率変数である。この例であれば、 z を $1, \dots, k$ までを取る離散の隠れ変数として、 $p(\mathbf{y}, z; \theta)$ は、

$$p(\mathbf{y}, z; \theta) = \sum_{i=1}^k \pi_i \delta_i(z) G(\mathbf{y}; \mu_i, \Sigma_i),$$

と書ける。ここで、 $\delta_i(z)$ は $z = i$ のときにのみ 1 を取る関数である。

4.3 Helmholtz マシン

脳の神経細胞には、神経細胞の集合から成る複数のモジュールがあり、それらが相互に結合している。このモジュール間の結合を用い、学習を行なうモデルとして Helmholtz マシンが提案された。Helmholtz マシンは生成モデル (generative model) と認識モデル (recognition model) の2つのモジュールからなるモデルである [4]。図7に示すのが Helmholtz マシンの模式図である。Helmholtz マシンは外部からの入力となる Visible variable と外部からは直接観測できない Hidden factor を持つ。これを脳のモジュール間の相互の結合に対応させ、visible variable を持つ細胞を低次の神経細胞、hidden factor を持つ細胞を高次の細胞と呼ぶこともある。

この2つの変数の組に対し、Helmholtz マシンでは2組のパラメータを与える。1つは visible variable を得たときに hidden factor の確率分布を与える認識モデル (recognition model) であり、も

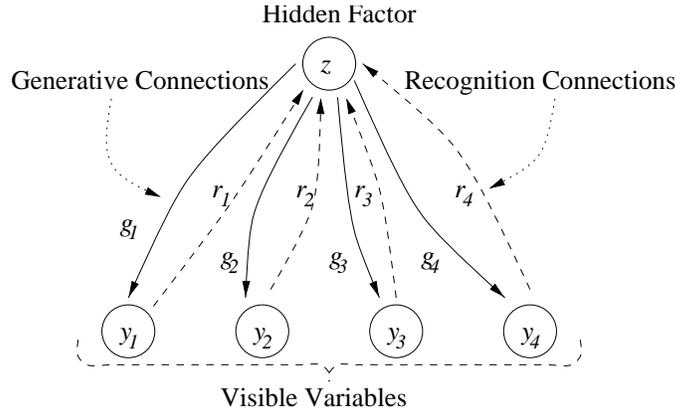


図 7: Helmholtz マシン

う 1 つは visible variable と hidden factor の両方の分布を与える 生成モデル (generative model) である。脳に対応させると、低次から高次への結合と高次から低次への結合が共存していることになる。

各々のモデルの形については Helmholtz マシン自体は規定していない。つまり、どのようなモデルを定義するかによって、Helmholtz マシンは様々な分布となる。例えば最も簡単な線型なモデルの場合、これは古典的な因子分析のモデルと一致する。また、混合正規分布や HMM (隠れマルコフモデル)、など隠れ変数を持つモデルは全てこの形で表現することができる。ここでは線型の場合のモデルについて定義をしておく [5]。

\mathbf{y} を n 次元の visible variable とし、 z を 1 次元の hidden factor とする。

- 生成モデル： \mathbf{y} を n 次元の信号が標準正規分布 $N(0, 1)$ にしたがう確率変数 z によって

$$\mathbf{y} = \mathbf{g}z + \boldsymbol{\varepsilon} \quad (12)$$

により生成されるとする。 $\boldsymbol{\varepsilon}$ は対角行列 $\Sigma = \text{diag}(\sigma_i^2)$ を分散行列とする正規分布 $N(0, \Sigma)$ にしたがう雑音である。このとき、 $p(\mathbf{y}, z; \mathbf{g}, \Sigma)$ は、

$$p(\mathbf{y}, z; \mathbf{g}, \Sigma) = G \left(\begin{pmatrix} z \\ \mathbf{y} \end{pmatrix}; \mathbf{0}, \begin{pmatrix} 1 & \mathbf{g}^T \\ \mathbf{g} & \Sigma \end{pmatrix} \right).$$

- 認識モデル：観測された信号 \mathbf{y} から対応する z が

$$z = \mathbf{r}^T \mathbf{y} + \delta \quad (13)$$

のように分布するとする。ただし δ は $N(0, s^2)$ にしたがう雑音である。認識モデルでは \mathbf{y} が観測されたときの z の条件付き分布を定義する。 $q(z|\mathbf{y}; \mathbf{r}, \sigma^2)$ とすると、

$$q(z|\mathbf{y}; \mathbf{r}, \sigma^2) = G(z; \mathbf{r}^T \mathbf{y}, \sigma^2).$$

以上のように、Helmholtz マシンは隠れ変数を持つモデルを表現する一つの手法であり、生成モデルと同時に認識モデルを同じ確率変数の組みに対し定義するところに特徴がある。

4.4 EM アルゴリズム

では、隠れ変数のあるモデルにおける最尤推定はどうであろう。ある確率変数 $X = \{Y, Z\}$ があり、その一部 Y のみが観測でき、残り Z は観測できない状況を考える。観測データ $\{y_1, y_2, \dots, y_T\}$

が得られたときに、確率モデル $p(y, z; \theta)$ のパラメタ θ を推定したいとする。

$$p(y; \theta) = \int p(y, z; \theta) dz$$

と定義されるが、この形は必ずしも単純ではなく、(10) 式を直接解くのは難しいことが多い。このような場合に用いられる手法の 1 つに EM アルゴリズムがある。

EM アルゴリズムは E-step (Expectation step) と M-step (Maximization step) の二つの部分からなり、これらを交互に繰り返してパラメタを更新することにより、最尤推定量あるいは尤度関数の極大点を得ることができる。

適当な初期値 θ_0 から始めて t 回更新した後のパラメタを θ_t として、E-step と M-step の具体的な手続きは以下のように定義される。

- **E-step**

次式で定義される $Q(\theta, \theta_t)$ を求める。

$$Q(\theta, \theta_t) = \frac{1}{T} \sum_{s=1}^T \left\{ \int p(z|y_s; \theta_t) \log p(y_s, z; \theta) dz \right\} \quad (14)$$

- **M-step**

$Q(\theta, \theta_t)$ を最大にする θ を求め、それを θ_{t+1} にする。

$$\theta_{t+1} = \underset{\theta}{\operatorname{argmax}} Q(\theta, \theta_t) \quad (15)$$

この結果得られた θ_t と θ_{t+1} との間には $\sum_s \log p(y_s; \theta_t) \leq \sum_s \log p(y_s; \theta_{t+1})$ という関係がある。

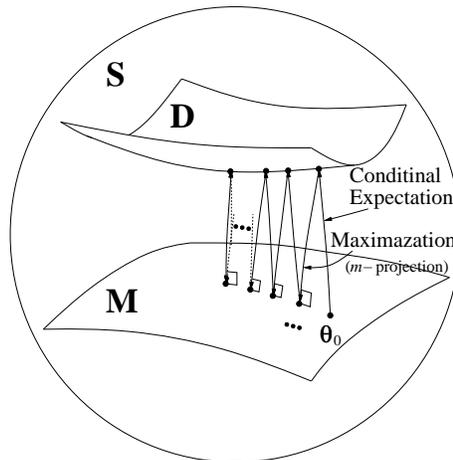


図 8: EM アルゴリズム

EM アルゴリズムを情報幾何的に解釈する。単純な問題では図 3 のように最尤推定は点から多様体への射影として捉えられる。一方、隠れ変数を持つモデルでは観測できる確率変数 y の確率分布の空間ではなく、確率変数 $x = \{y, z\}$ の確率分布の空間を考えた方が分り易い場合が多い。この空間を考えよう。

今、モデルのほうは前節と同様に 1 つの多様体 M を構成する。一方データのほうは y に関する経験分布 $q(y)$ しか与えない。このままでは $x = \{y, z\}$ の確率分布の空間中の点とはならないので、

z に関する任意の分布を付け加え D (図 8) という多様体を構成する (より詳しくは [3, 6] を参照されたい). EM アルゴリズムはこの 2 つの多様体の中のそれぞれの点で D と M とを最も近くする点を求めることに対応している. これを元に EM アルゴリズムを書き換えると甘利によって提案された *em* アルゴリズムとなる [6].

- *e-step* 多様体 D 上で $D(q(x; \eta), p(x; \theta_t))$ を最小にする η_{t+1} を求める.
- *m-step* 多様体 M 上で $D(q(x; \eta_{t+1}), p(x; \theta))$ を最小にする θ_{t+1} を求める.

5 まとめ

本稿ではエントロピーや情報量の考え方から、隠れ変数、その学習則までを簡単に説明をした. 脳の情報処理を考える際に、隠れ変数や情報量の考え方が役に立つと思う.

例えば、高次の情報と低次の情報を考えるときに、高次の処理では恐らく情報の集約が行なわれているが、そのとき、低次の神経細胞における情報のなかで、意味を持つ情報が保存されていることが望ましい. したがって、低次の神経活動を確率変数を X で表し、高次の活動を確率変数を Y で表すならば、 $I(X, Y)$ は小さい方が望ましいと考えられる. 確率変数、情報量の考え方は見通しを良くする一つの指針を与えると思う.

参考文献

- [1] 甘利 俊一. 情報理論. ダイヤモンド社, 1970.
- [2] 有本 卓. 確率・情報・エントロピー. 森北出版, 1980.
- [3] 甘利 俊一 and 長岡 浩司. 情報幾何の方法. 岩波講座 応用数学 [対象 12]. 岩波書店, 1993.
- [4] Peter Dayan, Geoffrey E. Hinton, and Radford M. Neal. The Helmholtz machine. *Neural Computation*, 7(5):889–904, 1995.
- [5] Shiro Ikeda, Shun-ichi Amari, and Hiroyuki Nakahara. Convergence of the wake-sleep algorithm. In Michael S. Kearns, Sara A. Solla, and David A. Cohn, editors, *Advances in Neural Information Processing Systems*, volume 11, pages 239–245. MIT Press, Cambridge, MA, 1999.
- [6] Shun-ichi Amari. Information geometry of the EM and em algorithms for neural networks. *Neural Networks*, 8(9):1379–1408, 1995.

予測と制御による階層的運動学習

春野雅彦

国際電気通信基礎技術研究所

1 はじめに

我々は日常生活において高度に複合的、階層的な運動を獲得して行動している。テニスや野球における個々の動作は要素運動が複雑に組み合わせられたものであるし、日常の会話でも文、単語、音素といった階層構造の中で発話を行っている。本講義では主としてこのような階層的運動がいかん獲得されるかについて計算論的に考察する。

階層的、複合的運動制御の学習に対しては2つのアプローチが考えられる。1つは単一のネットワークで問題全体を処理する方法で、他方は複雑な問題をより簡単な部分問題に分割し、各部分問題には少数の特化したネットワーク(エキスパート)を割り当てるモジュラネットワークによる方法である。前者と比較してモジュラネットワークによる方法は学習の局所性という顕著な特徴を持つ。すなわちモジュラネットワークでは適切に説明出来ない事例に対する学習の影響が適切なエキスパートに限定され、他のエキスパートには影響を及ぼさない。したがって早期の学習が可能となる。この学習の局所性ならびに神経生理学的に得られている機能局在やモジュラリティに関する知見から判断して脳における階層的運動学習のモデルとしてはモジュラネットワークの方が適切であると考えられる。

モジュラネットには適用する問題、部分問題への分割法、モジュールの構造及び使用する学習の手法によって様々な種類が考えられる。2章ではモジュラネットワークの例として Mixture of Experts [6] を挙げる。これは回帰モデルの一種でありゲーティングネットと呼ばれるネットワークで入力空間を分割し各部分空間に対してエキスパートネットを学習する。Mixture of Experts は工学的問題に用いられ成功を収めてはいるが、階層のあらゆる場所でゲーティングネットが入力そのものを分割する構造脳のモデルとして適切ではない。

3章では Multiple Paired Forward and Inverse Models (MPFIM) [8, 4] という運動制御の為に考案されたモデルの説明を行う。ここではゲーティングネットを用いず各エキスパートの予測の良さに基づき部分問題への分割が行われる。一般にモジュラネットではタスクを部分問題へ分割する必要があるがこれは池田先生の講義にある通り隠れ変数を持つ推定問題の一種である。MPFIM に簡単なマルコフ性のモジュール遷移を導入した2階層の場合を考察し EM アルゴリズム [1] による学習結果を示す。モデルから考案される行動実験についても簡単に触れる。4章では更に一般の階層的 MPFIM で得られた結果について簡単に紹介する。

2 これまでの研究

Mixture of Experts アーキテクチャ [6] は回帰モデルであり入力ベクトル x と出力ベクトル y の対 (x, y) を受取りそれらの間の関数関係を学習する。ゲーティングネット、 n 個のエキスパートネットは共に x を入力として受け取り、ゲーティングネットは i 番目のエキスパートネットの寄与率 g_i を、 i 番目のエキスパートネットは予測値 y_i を出力する。ゲーティングネットは各エキスパートネットに対応する n 個の出力を持ちその状態を s_i とすれば g_i は soft-max 関数 (1) で定義される。システム全体の最終的な予測値 y は (1) を用いた線形和 (2) で計算される。

$$g_i = \frac{e^{s_i}}{\sum_{j=1}^n e^{s_j}} \quad (1)$$

$$y = \sum_{i=1}^n g_i y_i \quad (2)$$

次にゲーティングネット、エキスパートネットにおける学習則を考える。ここでは (3) を目的関数として最急降下法による重みの更新を行う。(3) は各エキスパートの分離を良くする様に導入された便宜的な対数尤度であり y^* は教師信号、 σ_i は各エキスパートにガウスノイズを仮定した場合の標準偏差である。

$$\ln L = \ln \sum_{i=1}^n g_i e^{-\frac{1}{2\sigma_i^2} |y^* - y_i|^2} \quad (3)$$

ネットワークの学習は通常のバックプロパゲーションを用いて行う。はじめにゲーティングネットについて考える。(1) を用いてゲーティングネットの出力変数 s_i で対数尤度 (3) を偏微分すると (4) が得られる。ここで式中の h_i は (5) に示す i 番目のエキスパートが (x, y^*) を生成する事後確率である。(4) と (5) からゲーティングネットの学習はエキスパートネットの寄与率を事後確率に近付ける方向にすすむことが分かる。

$$\frac{\partial \ln L}{\partial s_i} = h_i - g_i \quad (4)$$

$$h_i = \frac{g_i e^{-\frac{1}{2\sigma_i^2} |y^* - y_i|^2}}{\sum_{j=1}^n g_j e^{-\frac{1}{2\sigma_j^2} |y^* - y_j|^2}} \quad (5)$$

同様にエキスパートネットの学習則を導出するためにエキスパートネットの出力変数 y_i で対数尤度 (3) を偏微分すると (6) が得られる。(6) から学習には事後確率 (5) で重み付けされた誤差が利用されることが分かる。

$$\frac{\partial \ln L}{\partial y_i} = h_i \frac{(y^* - y_i)}{\sigma_i^2} \quad (6)$$

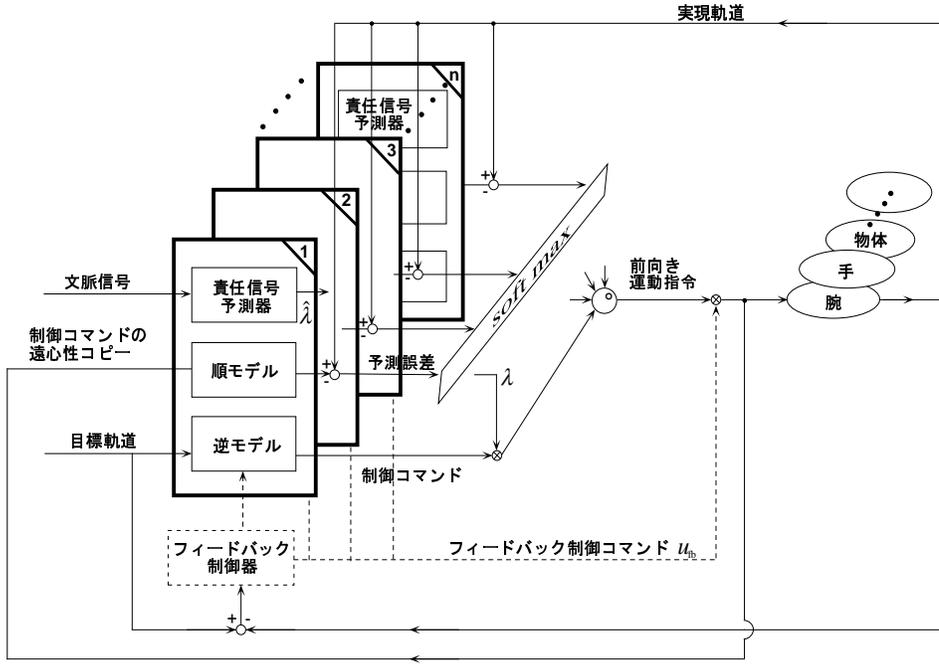


図 1: MPFIM アーキテクチャの概要

3 Multiple Paired Forward-Inverse Models

3.1 MPFIM の概要

本節では Multiple Paired Forward Inverse Model (MPFIM) [8, 4] と呼ばれるアーキテクチャについて説明する。図 1 に示す様に MPFIM の各モジュールは目標軌道から制御コマンドを計算する逆モデル、制御コマンドから実現軌道を予測する順モデル、視覚情報等の文脈情報からモジュールの適切さを予測する責任信号予測器の 3 つの部分から構成される。運動制御では各モジュールにおける制御コマンドの誤差を直接知ることが不可能であるため各制御コマンドから実現軌道を予測する順モデルを介する必要がある。順モデルの予測と目標軌道と比較することで現在の入力に適切なモジュールを選択し実際の制御及び学習を行うのである。

図 1 に示す n 個のモジュールを考える。このとき i 番目の順モデルは (7) の様に現在の状態 x_t 、制御コマンド u_t を入力として次時点の状態の推定値 \hat{x}_{t+1}^i を出力とする。ここで w_t^i はニューラルネットワーク ϕ の重みベクトルである。

$$\hat{x}_{t+1}^i = \phi(w_t^i, x_t, u_t) \quad (7)$$

順モデルによる予測を用いて各モジュールの適切さを示す責任信号 λ_t^i を (8) の soft-max 関数で定義する。 σ は雑音にガウス分布を仮定した場合の標準偏差である。責任信号は 0 と 1 の間の値を取り現在の状態を良く説明するモジュールほど大きな責任信号値を取る。(9) に示す様に責任信号は順モデルの学習にも利用される。責任信号値の大きなモジュールは多くの誤差を受け取り学習が進展する。責任信号を介した競合学習によってモジュールの分化が進むのである。

$$\lambda_t^i = \frac{e^{-|x_t - \hat{x}_t^i|^2 / 2\sigma^2}}{\sum_{j=1}^n e^{-|x_t - \hat{x}_t^j|^2 / 2\sigma^2}} \quad (8)$$

$$\Delta w_t^i = \epsilon \lambda_t^i \frac{d\phi_i}{dw_t^i}(x_t - \hat{x}_t^i) = \epsilon \frac{d\hat{x}_t^i}{dw_t^i} \lambda_t^i (x_t - \hat{x}_t^i) \quad (9)$$

次に逆モデルについて考える。(10)の様に*i*番目の逆モデルは目標状態 x_{t+1}^* と現在の状態 x_t を受取り制御コマンド u_t^i を出力する。ただし α_t^i はニューラルネットワーク ψ の重みベクトルである。

$$u_t^i = \psi(\alpha_t^i, x_{t+1}^*, x_t) \quad (10)$$

最終的な制御コマンド u_t は各逆モデルの出力を(11)の様に責任信号で重み付けることで得られる。また逆モデルの学習則は(12)で与えられ、順モデルの場合と全く同様に責任信号を用いて競合的に行われる。ただし逆モデルの場合には真の誤差信号 $u_t^* - u_t^i$ を知ることは出来ないためこれをフィードバック制御コマンド u_{fb} で近似している。

$$u_t = \sum_{i=1}^n \lambda_t^i u_t^i = \sum_{i=1}^n \lambda_t^i \psi(\alpha_t^i, x_{t+1}^*, x_t) \quad (11)$$

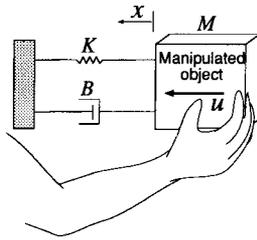
$$\Delta \alpha_t^i = \epsilon \lambda_t^i \frac{d\psi_i}{d\alpha_t^i}(u_t^* - u_t^i) = \epsilon \frac{du_t^i}{d\alpha_t^i} \lambda_t^i (u_t^* - u_t^i) \simeq \epsilon \frac{du_t^i}{d\alpha_t^i} \lambda_t^i u_{fb} \quad (12)$$

3.2 MPFIM の適用例

MPFIMを用いた時間的に変化する複数物体の制御シミュレーションについて述べる。図2に示す通り5秒毎に質量、粘性、弾性が α 、 β 、 γ と規則的に変化する物体を目標軌道(図3(b)の最下図)に沿って30秒間動かすことが課題である。この問題に対して3つのモジュールから成るMPFIMを構成した。このシミュレーションでは各順モデル、逆モデルとして線形ネットワークを用い、 σ は人手により最適値に設定した。線形ネットワークを用いたことで*i*番目の順モデル、逆モデルの重みベクトルからM,B,Kの推定値 M_i^F, B_i^F, K_i^F と M_i^I, B_i^I, K_i^I を計算することが出来る。図3(a)は学習に伴う M_i^F, B_i^F, K_i^F の推移を示しており、ランダムな初期値からはじまって α 、 β 、 γ へと収束していく。表1に $M_i^F, B_i^F, K_i^F, M_i^I, B_i^I, K_i^I$ の収束値を示すが、これらは α, β, γ の良い近似値となっている。表中の全ての場合において順モデルの推定値が逆モデルの推定値よりも優れている。これはまず順モデルから得られる責任信号を用いて逆モデルの学習が制御されるMPFIMの特徴を顕著に示している。

モジュール	M_i^F	B_i^F	K_i^F	M_i^I	B_i^I	K_i^I
1	1.0020	2.0080	8.0000	1.0711	2.0080	8.0000
2	5.0071	7.0040	4.0000	5.0102	6.9554	4.0089
3	8.0029	3.0010	0.9999	7.8675	3.0467	0.9527

表1: 順モデル、逆モデルで獲得された物体の物理特性



	M (Kg)	B ($\text{N m}^{-1} \text{ s}$)	K (N m^{-1})
α	1.0	2.0	8.0
β	5.0	7.0	4.0
γ	8.0	3.0	1.0

図 2: シミュレーション実験の概念図. 質量 M 、粘性 B 、弾性 K の異なる物体を目標軌道に沿って運動させる. 表は各物体の物理特性を示す.

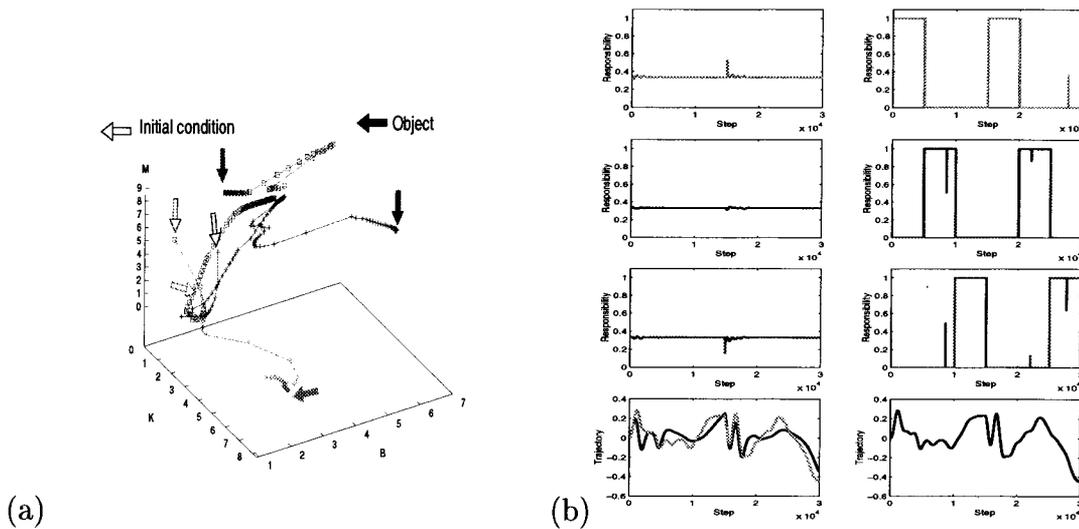


図 3: (a) ランダム値からはじめた 3 物体に対応する順モデルパラメータの学習. (b) 3 つのモジュールの責任信号と運動精度 (最下). 左は学習初期、右は学習終了時.

図 3(b) に学習初期 (左側) と学習終了後 (右側) の各モジュールの責任信号 (上 3 つ) と実現軌道 (最下) を示した。学習初期には各モジュールの分化が進んでおらず目標軌道と実現軌道の差が大きい。それに伴い各モジュールの責任信号も同様の値 (約 $1/3$) を取っている。これに対して学習終了時には各モジュールの責任信号が順に 1 を取りそれぞれが α 、 β 、 γ に対応していることが分かる。このときの実現軌道はほぼ目標軌道に一致している。

3.3 マルコフ的にモジュール遷移する MPFIM と EM アルゴリズムによる学習

ここでは Hidden Markov Model (HMM) [5] を用いてマルコフ的にモジュールが遷移する MPFIM を導入し EM アルゴリズムによる学習結果について述べる。上位階層として遷移ダイナミクスを導入した最も簡単な階層モデルと考えられる。HMM では各モジュール (状態) 間の遷移確率からなる遷移確率行列 P を定義する。 a_{ij} はモジュール i からモジュール j に遷移する確率で時間によらず一定であるものとする。

$$P = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

HMM では時刻 t におけるモジュール S_t は N 個のモジュールのうちどれかを取りその間の状態遷移は状態遷移行列 P で規定される。加えて各モジュールにおいて順モデルはそれぞれの出力を行う。 i 番目のモジュールの順モデルが持つ線形パラメータを W_i 、標準偏差を σ_i とし、入力を X_i とするとこのモジュールが観測データ Y を生成する尤度 $L_i(Y|W_i, \sigma_i)$ はガウスノイズの仮定のもとで以下ようになる。

$$L_i(Y|W_i, \sigma_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left\{-\frac{1}{2\sigma_i^2}(Y-X_iW_i)'(Y-X_iW_i)\right\} \quad (13)$$

状態遷移行列も考慮しシステム全体の尤度の期待値を計算すると (14) のようになる。 θ は順モデルのパラメータ (W_i, σ_i) を示す。また、 ω 、 s_t はそれぞれ可能なモジュール列の全体、ならびに時刻 t に選択されるモジュールである。

$$L(Y|\theta) = \sum_{\omega} \prod_{t=0}^{T-1} a_{s_{t-1}s_t} L_{s_t}(Y(t)|W_{s_t}, \sigma_{s_t}) \quad (14)$$

詳細は省略するが (14) を最大化するパラメータ P 、 W_i 、 σ_i は EM アルゴリズムの 1 種である Baum-Welch アルゴリズム [5] を用いて反復的に解くことが出来る。また逆モデルのパラメータは通常のフィードバック誤差学習によって得られる。

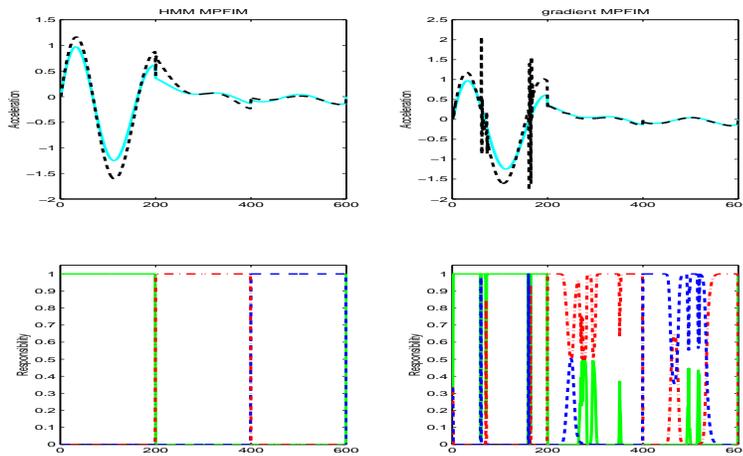


図 4: HMM と通常の学習法における実現された加速度とモジュールの比較

図 4は 3.2 節のシミュレーション課題に HMM (左側) と 3.1 節の式 (9) (右側) を適用した結果を示す [3]。このシミュレーションでは HMM のパラメータ σ_i は学習アルゴリズムにより自動的に推定しているが、式 (9) で用いる σ は人手により最適な一定値に設定している。上側のパネルにおいて実線は目標加速度、点線は実現された

加速度を示し、下側のパネルは各モジュールの責任信号の変化を示す。HMM ではモジュールの切り替えが適切に行われているのに対し、(9)の方法では頻りにモジュールの切り替えが起こっている。これに伴ってHMMでは加速度の誤差も小さくなっている。これらの差はHMMでは状態遷移確率を考慮してモジュールの切り替えを安定化させているために生じていると考えられる。

更にHMMの効果を検討するために様々な切替え数、初期状態で両者の比較を行った。表2に様々な間隔でモジュールを切り替えた場合の両者の誤差の比を示す。各ネットワークの初期値として真の値を中心に20%、40%のノイズを加え50回の試行の平均値を示した。この結果から初期値が真の値から遠く、モジュール切替えの頻度が低い場合にはHMMの精度が顕著に高くなることが分かる。このようにモジュラネットの学習にEMアルゴリズムを適用することで状態遷移行列 P や競合学習のパラメータ σ_i などを自動的に推定出来るだけでなく、安定して高い精度を得られることが分かる。

Initial conditions s.d.	Switching period				
	5	25	50	100	200
-					
20%	0.998	0.887	1.031	0.636	0.443
40%	0.855	0.676	1.147	0.598	0.322

表 2: 2種類のノイズに対する加速度誤差の比: HMM-MPFIM/gradient-MPFIM

実際の脳においてもモジュール遷移のダイナミックスのトップダウン情報と実際の誤差に基づく情報の両者のせめぎあいでの制御が行われていると考えられる。したがってHMMに基づくモジュール遷移を持つ実験課題を設定し

- 受動的に制御を行う
- 状態遷移は見せて探索的に制御を行う
- 状態遷移も見せずに探索的に制御を行う

を比較することでパラメータの推定、状態の予測に関する脳部位を特定出来るものと考えられる。

4 階層的MPFIM

これまで述べたMPFIMを更に一般の階層的な場合に拡張した階層的MPFIMについて述べる。図5(a)に示すとおり各階層はMPFIMから構成されている。一番下位の階層では制御(逆)モデルが制御コマンドを出力し予測(順)モデルはそのモジュールの良さを予測する。これに対して上位のモジュールでは予測モデルの役割は同じであるが制御モデルが自分の下の階層の各モジュールに対する事前確率分布を出力する。つまり直接制御コマンドを出すのではなく抽象的な形で下位のモジュールを制御している。

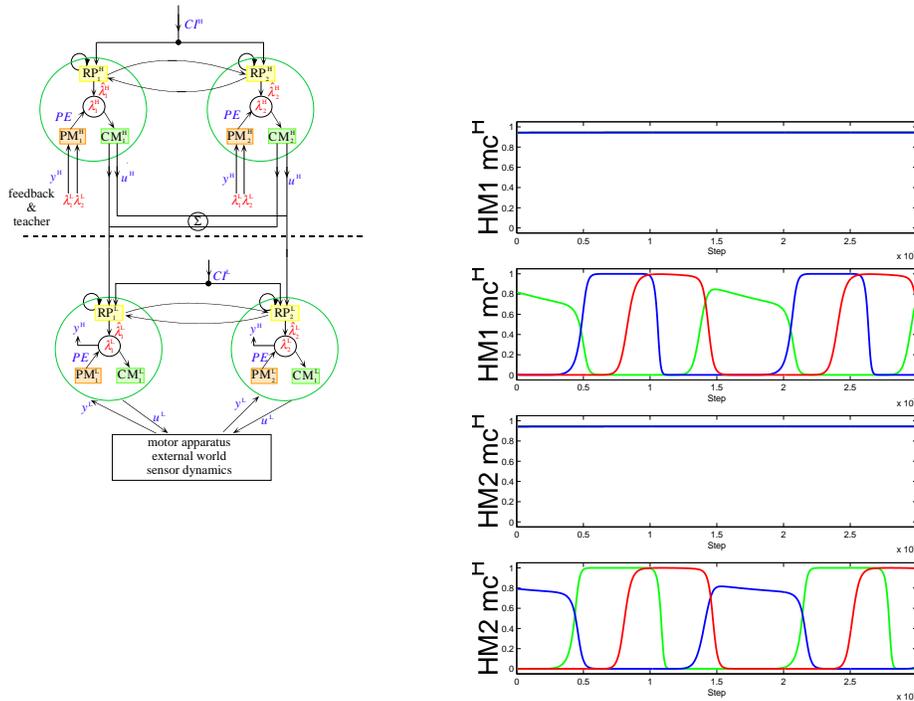


図 5: (a) 一般化された階層的 MPFIM の概念図、(b) 2つのシーケンスを学習する前後の上位 2 モジュールの出力。上から学習前、学習後の順に並んでいる

下位モジュールを 3 つ、上位モジュールを 2 つ持つ階層的 MPFIM を用意し、2.2 節で述べたシミュレーションにおいて 2 種類の系列: 系列 A: $\alpha \rightarrow \beta \rightarrow \gamma$ and sequence 系列 B: $\beta \rightarrow \alpha \rightarrow \gamma$ を学習させた。下位の 3 モジュールは図 3(b) と同様に収束した。上位 2 モジュールの出力を図 5(b) に示す。学習前は全く予測を行えず同じ確率を出力しているが学習後は各モジュールが 2 つの系列に対応して下位モジュールを選択していることが分かる。このように運動系列に固有の反応をするニューロンは猿の SMA や PreSMA でも見つかっており [7]、運動制御の階層性、双方向処理を考える上で興味深い。

5 おわりに

階層的制御モデルとして MPFIM を導入し、学習によって様々な階層性が獲得されることを述べた。高等生物の脳における運動制御ではモジュール的、階層的処理が本質的な役割を果たしており、ここで述べた様な手法は脳の情報処理に関する様々なレベルでのモデル化や実験のデザインに寄与し得る [2]。

参考文献

- [1] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum-likelihood from incomplete data via the EM algorithm. *J. Royal Stat. Soc. B*, 39:1–38, 1977.
- [2] Z. Ghahramani and D.M. Wolpert. Modular decomposition in visuo-motor learning. *Nature*, 386:392–395, 1997.
- [3] M. Haruno, D.M. Wolpert, and M. Kawato. Multiple paired forward-inverse models for sensorimotor learning and control. *submitted for publication*, 1999.
- [4] M. Haruno, D.M. Wolpert, and M. Kawato. Multiple paired forward-inverse models for human motor learning and control. In M. Kearns and S. Solla, editors, *Advances in Neural Information Processing Systems 11*, pages 31–37. MIT Press, 1999.
- [5] X.D. Huang, Y. Ariki, and M.A. Jack. *Hidden Markov Models for Speech Recognition*. Edinburgh University Press, 1990.
- [6] R. Jacobs, M. Jordan, S. Nowlan, and G. Hinton. Adaptive mixture of local experts. *Neural Computation*, 3:79–87, 1991.
- [7] J. Tanji and K. Shima. Role for supplementary motor area cells in planning several movements ahead. *Nature*, 371(29):413–416, 1994.
- [8] D.M. Wolpert and M. Kawato. Multiple paired forward and inverse models for motor control. *Neural Networks*, 11:1317–1329, 1998.

前運動野・一次運動野における運動指令の座標系に見られる階層性について：空間座標から筋肉座標へ

寛 慎治（科学技術振興事業団・東京医科歯科大学）

1. はじめに

目標到達運動（goal-directed movement, 例：コップに手を伸ばす運動）の脳内過程は、外部座標系で表現された目標の位置が、適切な筋活動のパターンへと変換される「トップダウンの階層的な座標変換」とする見方が一般的である。脳内で運動指令がどのようなニューロン活動で表現されているにせよ、「手 腕」というマニピュレーターの目標への到達を保証する定量的なものでなければならない。従って、何らかの基準点を持った座標系によって記述されているはずである。残念ながら（セミナー参加者の皆さんには幸いなことであるが）この運動指令を直接解読し「座標変換」のアルゴリズムを理解することが近い将来可能になる見込みは著しく低い。では、問題をよりコンパクトな形にスケールダウンしてアルゴリズムの骨格を概観する方法は無いのか？一つの戦略は、「運動指令をコードするニューロン活動の座標系を神経回路の接続に沿って同定し、座標系の遷移から逆にアルゴリズムを推定する」というものである（この考え方の根底には、脳内の情報表現[符号化の方法(座標系)]は、それが処理されるアルゴリズムと密接な関連があるという期待が込められている）。目標到達運動は、座標変換のはじめ（目標の位置）とおわり（筋活動）が明確に定義できるので、この戦略を適用して中間表現を推定するのに適していると考えられる。ところがこの様に見る簡単な研究が過去に行われたことはない。

なぜだろうか？

2. これまでの研究

運動指令が脳の中でどの様にコードされているかを明らかにする試みは、Evarts の記念碑的な実験まで遡る（Evarts, 1968）。Evarts は運動課題を訓練したサル的一次運動野（M1）から、運動実行中のニューロン活動を記録することに初めて成功した。彼の結論は、運動野の出力ニューロンの活動が、主として末梢の関節トルクや、筋活動に相関するというものであった。これはM1出力ニューロンの主要な座標系が、末梢の関節或いは筋肉等の身体座標系（一般的には内部座標系という）であることを示唆する。ところが80年代に入り、Georgopoulosら（1982）は、同じM1ニューロンが、関節トルクや筋活動のような身体座標系ではなく、より抽象的な外部座標系で運動の方向をコードしているという対照的な主張を展開し、両者の見解が対立した。一方、この対立とは独立に70-80年代から、ロボティクスにおける「フレーム」（ここでは一応「座標系」と和訳する）の概念（注）が

工学系の研究者によって目標到達運動の定式化に取り入れられた（総説として Soechting & Flanders 1992 を参照）。ここでポイントになるのは外部座標系と身体座標系を明確に区別することである。身体座標系での運動指令は関節や筋肉の状態に直接大きな影響を受けるのに対して、外部座標系における運動指令はこれらの身体の状態に無関係であるので、両者の異質性とそれを区別する重要性は理解していただけたと思う。Evarts も Georgopoulos も、異なる座標系の概念を意識していたにも関わらず、身体座標系 vs 外部座標系の対比では十分な分離が得られていなかった。歴史的に見て、これはやむをえないことであった。なぜなら、身体座標系 vs 外部座標系の対比に焦点を当てたその後の実験（Kalaska ら, 1989; Alexander と Crutcher, 1990; Scott & Kalaska, 1997）においてさえ、外部座標系と身体座標系の分離という最も重要な点が、理論的にも（実験のデザインの欠陥により）不十分であり、肝心の座標系の「同定」には使えなかった。前に述べた一見「簡単な」実験が、実は簡単ではなかったのである。

直面する問題は明確である。どの様にして複数の座標系、特に外部座標系と身体座標系を分離できるような実験系をデザインするかに尽きる。

（注：ここで使っている「フレーム」は運動を記述する座標系の固定（基準）点の位置を指すが、その基準点での座標系の姿勢（例えば人差し指の先端に $x y z$ 座標系を設定する場合に x 軸をどちらに向けるとか）には関心がない。その意味でロボティクスの「フレーム」よりも緩い概念である。また、具体的な座標系のタイプ（ $x y z$ であるとか極座標であるとか、 \dots ）も想定しない。独立して動くものには別々のフレームを設定でき、一緒に動く場合はフレームを「共有」しているという。例 1：腕時計が手首に固定されているとき、腕時計のフレームは手首のフレームに固定されている。あるいは腕時計と手首が一つのフレームを共有している。例 2：外部空間の空間座標系のフレームは身体の中のどの部位のフレームとも独立している。）

3. 最近の研究

Takei ら（Takei ら, 1999）による実験は、この問題に対する一つの解答である。Takei らも、Evarts の実験と同様にサルをトレーニングして、マニピュランダムを使った手首（wrist）の単関節の運動を習得させた。サルは眼前の CRT 上のカーソルを、それに連動したマニピュランダムを動かしてコントロールする。イメージとしてはテレビゲームのジョイスティックをマニピュランダムで置き換えていただければよい。この実験では、伸展 屈曲の 1 自由度に加えて、それに垂直な軸のもう 1 自由度を加えた 2 次元で、中心から上下左右とその中間の 8 方向への運動を行わせる（Takei ら, 1999 の図 1B を参照）。もう一つの、そして最大の特徴は、この 2 次元の手首の運動を、3 つの異なる前腕の姿勢で行わせる点にある（図 1A）。図の一番左は、Prone ポジションで、ちょうど椅子の肘掛けに前腕を乗せた姿勢である。それに対して右は Supine ポジションで（血圧を測るときの姿勢）

左の Prone ポジションから前腕を 180° 時計回りにまわした状態、そして真ん中は両者の中間である。このタスクは外部座標系での運動の方向と手首の関節座標系（身体座標系の一つ）での運動の方向が分離できるようにデザインされている。例えば、Prone では手首の「伸展（関節座標系での記述）」は「上向き（外部座標系での記述）」の運動であるが、前腕を 180° 回転して Supine にすると今度は、同じ手首の「伸展（関節）」が「下向き（外部）」の運動になる。つまり関節座標系では同じ「伸展」運動が、前腕の姿勢を変えたことにより外部座標系では異なる運動になる。この様なタスクを十分に訓練されたサルで、タスク実行中の 27 個の前腕 上腕 肩の筋肉から筋電図を記録したところ、手首の運動に関連した活動を示したのはたった 7 個の前腕の筋肉だけであった。これら 7 個の筋活動全てが、運動の方向に関していわゆる cosine tuning (Georgopoulos ら, 1982 を参照) を示した。従って、各筋について、各ポジションでの活動のピークの方向 (preferred direction, PD) を cosine fitting で計算できる。図 1C にはそのようにして計算された 3 個の筋の PD が矢印で示してある。図 1C の左の円に示した Prone のデータを見ると ECRB という筋の PD はおおよそ 11 時の方向であることがわかる。Prone から前腕を時計回りに 180(回転した Supine (右) では ECRB の PD も同じ時計回りに回転し、2 時のあたりに移動している。これは少しおかしい(?)。前腕が 180(回転しているのに PD は高々 90(程度しか回転していないことになる。しかもこの「中途半端な回転」を示しているのは ECRB だけではない。図 1C に示した他の 2 つの筋も同様である (実は 7 個の筋全てに共通!)。そういう眼で図 1C の 3 つのグラフを眺めると、異なる 3 個 (実は 7 個全て) の筋の PD が星座の様に相対的な位置関係を保ったまま「一緒に」回転している。これは「座標系 (フレーム)」という観点に立てば、これらの異なる筋活動が共通の座標系」に乗っていることを示唆する。しかも「中途半端な回転」が幸いして、この座標系は外部座標系とも手首の関節座標系とも明瞭に区別できる (図 1D を参照)。従って理論的にも実用上も、この座標系を「(この実験系における) 筋肉座標系」と呼ぶことに同意していただけたらと思う。改めて言うまでもないが、筋肉と関節は機能的に極めて密接な関係にあるので、筋活動と関節の座標系がこれほど明瞭に分離できたことは、驚きであった。

以上まとめると、この実験は、Prone, Middle, Supine, の 3 つのポジションで 8 方向の運動を行わせ、PD の回転量を計算することにより、「外部座標系」、「筋肉座標系」、「関節座標系」、そして「それ以外の座標系」のどれに属するかを判別できることが明らかになった。ここで強調しておかなければならないことは、この解析方法は方向選択性を持った (特定の 1 つの方向にピークを持つ) 活動であれば、筋活動のみならずニューロン活動の座標系の同定にもそのまま使えるという適用範囲の広さである。そこで次のステップとして、このタスクを、まず M1 ニューロン活動の座標系の分析に試用し、その効果を確認した (Kakei ら, 1999)。次いで前運動野のニューロン活動の解析に拡張して (投稿準備中)、先に述べた階層的座標変換の分析への有効性を確認した。当日は、これらのニューロン活動のデータを中心に述べる予定である。また、そこから示唆された前運動野と M1 を中心

とする階層的なネットワークのモデルを提示し、このモデルと春野らの「複数の内部モデルからの選択」のモデルとの類似性を議論したい。最後に、これらの皮質領野間に見られる強力な双方向性結合の意義を座標系の階層性という視点から検討する予定である。

4 . おわりに

以上のように、この新しい実験によって「運動指令をコードするニューロン活動の座標系を神経回路網の接続に沿って同定し、座標系の遷移から逆にアルゴリズムを推定する」という戦略を実行に移す環境が整った。手首の運動に関連する活動を示す記録可能なニューロンであれば、この実験系によりその座標系の同定が可能である。頭頂連合野、補足運動野、体性感覚野、小脳、大脳基底核は、この方法で解析されるべき対象であろう。特に、双方向性の結合が少ない、前向きの神経結合が多いところでは、入力と出力の関係が明確であり、データの解釈が比較的容易であると考えられる。そのような条件を持つ脳内の部位として、大脳小脳連関と、大脳と基底核を結ぶループのシステムは、この実験系での解析が特に有望な対象であると考えられる。中でも大脳小脳連関は、伊藤正男や川人光男らによって（例えば、川人・銅谷・春野「科学」2000年7月号を参照）運動制御を初めとする様々な高次脳機能の「内部モデル」の座であることが提案され、実験的にも（例えば Imazumiら, Nature 403:192, 2000）それを裏付けるデータが集積されつつある、最も興味深い対象である。もちろんこの実験系はいわゆる高次脳機能を直接扱えるものではないが、大脳小脳連関における神経回路の特異的と言ってもよい並列性から考えて、運動機能について得られた回路のアルゴリズムは、他の高次機能にも相当部分適用できる可能性が高いと期待している。

5 . 参考文献:目を通しておいてほしいもののみリストアップしました。

Evarts, EV. Journal of Neurophysiology, 31:14 (1968). 記念碑的論文。

Georgopoulos, AP. Et al. Journal of Neuroscience, 2:1527 (1982). 記念碑的論文。

Soechting, JF., Flanders, M. Annual Review of Neuroscience, 15:167 (1992). フレームの概念が平易にかつ詳しく書かれています。

Takei, S., Hoffman, DS., Strick, PL. Science, 285:2136 (1999).

NISS2000 ファカルティ 一覧

校長

塚田 稔 (玉川大学) tsukada@eng.tamagawa.ac.jp

プロデューサー

川人 光男 (ATR人間情報通信研究所) kawato@hip.atr.co.jp

ディレクター

五味 裕章 (NTTコミュニケーション科学基礎研究所) gomi@idea.brl.ntt.co.jp

阪口 豊 (電気通信大学) sakaguchi@is.uec.ac.jp

銅谷 賢治 (国際電気通信基礎技術研究所) doya@erato.atr.co.jp

ファカルティ

Rajesh P. N. Rao (Salk Institute) rao@salk.edu

Watanabe Takeo (Boston University) takeo@bu.edu

池田 思朗 (科学技術振興事業団) shiro@brain.riken.go.jp

石井 信 (奈良先端科学技術大学院大学) ishii@is.aist-nara.ac.jp

今水 寛 (科学技術振興事業団) imamizu@erato.atr.co.jp

岡田 真人 (科学技術振興事業団) masato@erato.atr.co.jp

筧 慎治 (科学技術振興事業団・東京医科歯科大学) skakei.phy1@med.tmd.ac.jp

柏野 牧夫 (NTTコミュニケーション科学基礎研究所) kashino@av-hp.brl.ntt.co.jp

片山 正純 (豊橋技術科学大学) katayama@ics.tut.ac.jp

北澤 茂 (電子技術総合研究所) kitazawa@etl.go.jp

小池 康晴 (東京工業大学 精密工学研究所) koike@pi.titech.ac.jp

佐藤 雅昭 (国際電気通信基礎技術研究所) masaaki@hip.atr.co.jp

鮫島 和行 (科学技術振興事業団) samejima@erato.atr.co.jp

中原 裕之 (理化学研究所) hiro@brain.riken.go.jp

西田 眞也 (NTTコミュニケーション科学基礎研究所) nishida@apollo3.brl.ntt.co.jp

春野 雅彦 (国際電気通信基礎技術研究所) mharuno@hip.atr.co.jp

彦坂 興秀 (順天堂大学) hikosaka@med.juntendo.ac.jp

事務局

片山 直美 (科学技術振興事業団) katayama@erato.atr.co.jp

神経情報科学サマースクール NISS2000

2000年8月8日 - 13日, 神奈川県葉山町, 湘南国際村

主催: 日本神経回路学会

E-mail: jnns@jnns.inf.eng.tamagawa.ac.jp

<http://jnns.inf.eng.tamagawa.ac.jp/>

事務局: 619-0288 京都府相楽郡精華町光台 2-2

科学技術振興事業団 CREST 内

E-mail: niss2000@meta.rad.atr.co.jp

<http://meta.rad.atr.co.jp/niss2000/index.html>