

博士論文

行動予測に基づくヒト高次視覚情報処理の研究

鹿内 学

2010年3月17日

奈良先端科学技術大学院大学
情報科学研究科 情報生命科学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
博士(理学) 授与の要件として提出した博士論文である。

鹿内 学

審査委員：

石井 信 教授	(主指導教員)
湊 小太郎 教授	(副指導教員)
柴田 智広 准教授	(副指導教員)

行動予測に基づくヒト高次視覚情報処理の研究*

鹿内 学

内容梗概

脳の情報処理を理解するためには、次に挙げる3つの水準について理解する必要がある。1つは、脳が計算する目的や最適性を議論する計算理論の水準であり、次に計算理論を実現するために脳が採用するアルゴリズム・情報表現の水準であり、最後に、アルゴリズムを実装する脳の物理的特性であるハードウェアの水準である。本研究では、実験課題中のヒトの行動を予測することにより、脳情報処理の計算理論とアルゴリズム・情報表現を解明する研究を行った。

まず、ヒトの視覚眼球運動系、特にサッカードと呼ばれる眼球運動系における予測について検証した。計算論的に考えて予測は脳の情報処理において不可欠な要素だが、サッカードの系においてダイナミクスを用いて予測していることを実験的に示した研究はこれまでになかった。本研究では、実験課題を数理的に定式化し、最適性と実現のための論理を考え、つまり、計算理論の水準を考慮した行動実験を設計した。その上で、行動実験におけるヒトの行動を、数理モデルを用いて予測し、ダイナミクスを持つ視標運動に対してヒトがダイナミクスを用いて予測していることを解明した。また、ヒトは異なる視標運動のダイナミクスにあわせて適応的に予測していることが示唆された。

次に、fMRI計測実験を行い、高次元の視覚的情報表現について、脳情報の再構成と行動予測による検証を行った。顔知覚に関わる下側頭葉の情報表現の数理モデルを用いて、顔知覚課題中の下側頭葉のfMRI脳活動データからヒトが知覚した顔情報を画像として再構成した。さらに、再構成した顔知覚の情報をもとに顔識別行動を数理モデルにより予測し、ヒトの識別行動と比較した。その結果、少なくとも再構成できた程度の情報が下側頭葉の脳活動に存在することを示し、再構成精度を行動予測誤差により評価した。また、脳活動情報を画像として再構成したことにより、下側頭葉における高次元の視覚的情報表現を、感覚的に理解することを可能にした。

キーワード

脳、計算理論、情報表現・アルゴリズム、行動予測、眼球運動計測、fMRI

*奈良先端科学技術大学院大学 情報科学研究科 情報生命科学専攻 博士論文, NAIST-IS-DD0561016, 2010年3月17日.

Studies with the prediction of behavior for higher visual information processing in human brain*

Manabu Shikauchi

Abstract

For understanding of brain information processing, we must explain that in three levels; “computational theory”, “algorithm and representation” and “hardware implement”. This thesis aims at investigating mainly for understanding of “computational theory” and “algorithm and representation” in human brain, by a manner of predictions of human behavior.

In the first study, through recording of eye movements, I investigate an information processing to follow a moving visual target by eye movements, especially in saccade system. This study aims at investigating whether humans can predict aperiodic visual target sequences by a dynamics-based approach. For an explicitly considered computational theory of experimental tasks, I defined visual target sequences mathematically, and considered a logic of optimality prediction in target sequences. These results suggested that the prediction of a next target position the AR sequence was more successful than that by the random guess, and that participants attempted to predict adaptively the next target position.

In the second study, I investigated the representation of perceived visually face images with analysis of reconstructing a face information from the fMRI signals and the prediction of human’s face discrimination behavior. I formulated the representation of face perception and human discrimination behavior. The representation was formulated into face space model based on principal component analysis (PCA), and face discrimination behaviors were explained by the signal detection theory. Reconstructed information and predictions of human behavior demonstrated that there is information of face perception and how to understand intuitively the representation of high dimensional information in brain.

Keywords:

Brain, Computational theory, Algorithm & representation, Prediction of behavior, Eye-movements, fMRI

*Doctoral Dissertation, Department of Bioinformatics and Genomics, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD0561016, March 17, 2010.

目次

第1部 研究思想	2
序論	3
本論文の構成	6
第1章 研究対象と方法論	7
1. 視覚系と高次情報処理	7
2. 行動予測	8
第2部 高次視覚情報処理の研究	10
眼球運動計測を用いた系列予測の研究	11
第2章 系列予測の研究背景と目的	12
1. 予測のための脳情報処理	12
2. 本研究の目的	13
第3章 実験方法	15
1. 系列予測サッカード課題	15
2. 実験手順	15
3. 実験参加者	18
4. 実験設定の詳細	18
4.1 視標提示	18
4.2 実験機器と解析装置	18
4.3 計測と前処理	19
第4章 視標系列の計算理論	21
1. 視標系列	21
1.1 自己回帰過程	21
1.2 乱数系列	22

2.	視標系列の分析	23
2.1	予測誤差：最適予測時の予測誤差	23
2.2	AR 次数	23
2.3	AR パラメタの推定	24
2.4	最適予測の場合の行動	26
3.	計算理論のまとめ	31
第 5 章	系列予測行動結果	32
1.	予測誤差	32
1.1	対照実験との比較	32
1.2	学習について	32
2.	実験参加者の予測行動系列解析	37
2.1	実験参加者の予測行動の AR 次数	37
2.2	実験参加者の予測行動の AR パラメータ	37
3.	行動予測	38
第 6 章	系列予測についての議論	40
1.	ダイナミクスをもつ系列の予測	40
2.	ヒトの予測行動の特性	40
3.	行動予測の妥当性	41
	fMRI 計測を用いた顔知覚の研究	43
第 7 章	顔知覚の研究背景と目的	44
1.	顔知覚	44
1.1	顔知覚の計算理論；コンピュータビジョンの先行研究	44
1.2	脳の顔知覚の情報表現：生理学，心理学の先行研究	45
2.	本研究の目的	45
第 8 章	実験方法	46
1.	顔識別課題	46
2.	顔知覚空間	46
3.	顔識別モデル	48
4.	実験設定の詳細	48
4.1	原顔画像	48
4.2	平均顔画像の作成	50
4.3	変形顔の作成	52
4.4	実験で用いた顔画像	53

第 9 章 顔識別行動結果	57
第 10 章 脳活動解析	59
1. fMRI 研究	59
1.1 fMRI 計測	59
1.2 これまでの fMRI 研究	60
1.3 fMRI デコーディング研究	60
2. 脳活動解析の目的	62
3. fMR 実験	62
3.1 fMRI 計測設定	62
3.2 顔領域同定課題	62
4. 顔の距離に相関する脳活動	63
5. 顔知覚座標の推定	63
第 11 章 顔知覚についての議論	66
第 3 部 全体の議論	67
第 12 章 行動予測について	68
1. 予測の有用性について	68
1.1 生成モデル	68
1.2 一致性がないことへの弁明	69
1.3 予測する手続きによる理解の確認	71
2. 「行動」の予測	71
補足	74
全タスク・全実験参加者の予測系列	74
全タスク・全実験参加者の行動予測	78
全実験参加者の顔識別行動	84
謝辞	90
参考文献	94

目 次

3.1	系列予測サッカード課題・1 試行	16
3.2	系列予測サッカード課題・実験進行	17
3.3	眼球運動計測装置	19
4.1	AR 過程の系列	22
4.2	視標系列の BIC 値	24
4.3	逐次最小二乗法による計算機実験・AR 系列 (忘却係数: 1.00)	27
4.4	逐次最小二乗法による計算機実験・乱数系列	28
4.5	逐次最小二乗法による計算機実験・AR 系列 (忘却係数: 0.97)	29
4.6	視標系列の周波数分析	30
5.1	実験参加者の予測誤差: AR タスクの前 150 試行と後 150 試行の比較	33
5.2	実験参加者の予測系列の 1 例	34
5.3	実験参加者の予測誤差・AR タスクと乱数タスク比較	35
5.4	実験参加者の AR 次数推定・AR タスク	35
5.5	実験参加者の AR 次数推定・乱数タスク	36
5.6	実験参加者行動の予測の 1 例, AR_1 タスク, NN	39
8.1	顔識別課題	47
8.2	顔識別モデル	49
8.3	20 の原顔画像	50
8.4	顔画像の対応点	51
8.5	顔知覚空間における原点となる平均顔作成アルゴリズム	52
8.6	変形顔作成および顔知覚再構成アルゴリズム	53
8.7	実験で用いた顔画像	54
8.8	顔画像の顔知覚空間上での距離	55
8.9	顔知覚空間上における各顔の位置	56
9.1	顔識別行動の一例, 実験参加者 IN	58
10.1	機能的顔領野同定課題	64
10.2	顔の距離に相関する脳部位	65

12.1 視標系列の AIC, BIC による次数同定比較	70
12.2 実験参加者の予測系列・ AR_1 タスク	74
12.3 実験参加者の予測系列・ AR_{21} タスク	75
12.4 実験参加者の予測系列・ AR_{22} タスク	75
12.5 実験参加者の予測系列・ rnd_1 タスク	76
12.6 実験参加者の予測系列・ rnd_{21} タスク	76
12.7 実験参加者の予測系列・ rnd_{22} タスク	77
12.8 実験参加者行動の予測・ AR_1 タスク, KT	78
12.9 実験参加者行動の予測・ AR_1 タスク, MO	79
12.10 実験参加者行動の予測・ AR_1 タスク, NN	79
12.11 実験参加者行動の予測・ AR_1 タスク, TY	80
12.12 実験参加者行動の予測・ AR_{21} タスク, KY	80
12.13 実験参加者行動の予測・ AR_{21} タスク, MO	81
12.14 実験参加者行動の予測・ AR_{21} タスク, MO	81
12.15 実験参加者行動の予測・ AR_{22} タスク, KY	82
12.16 実験参加者行動の予測・ AR_{22} タスク, MO	82
12.17 実験参加者行動の予測・ AR_{22} タスク, NN	83
12.18 顔識別行動, AI	84
12.19 顔識別行動, IN	85
12.20 顔識別行動, MF	85
12.21 顔識別行動, MO	86
12.22 顔識別行動, MS	86
12.23 顔識別行動, NH	87
12.24 顔識別行動, SH	87
12.25 顔識別行動, TT	88
12.26 顔識別行動, YS	88
12.27 顔識別行動, YT	89

表 目 次

4.1	実験で用いた系列の特性	26
5.1	実験参加者の予測行動の AR パラメータ	37
5.2	特異的な戦略の有無	38
10.1	各手法における顔座標推定誤差	63

第1部 研究思想

序論

脳科学¹ は、生物学、生理学、心理学、物理学、工学、情報学、文化人類学や哲学などの学問からなる学際的学問として研究されている。学際的出る1つの理由として、脳の情報処理が大小様々な機能の階層性から実現されていることが関連する。脳の機能単位を考えた場合、分子や神経細胞の活動で実現される機能や、複数の神経細胞の結合で構成される局所回路で実現される機能、脳領野や領野間の結合を考えた脳全体での機能、などいくつかの機能単位があり、それぞれの機能単位について様々な方法論によって研究がなされている。例えば、分子から局所回路、脳領野の機能が、分子生物学的方法や電気生理学的方法によって調べられてきた。また、認知科学など心理学が対象としてきた機能は、例えば、知覚や記憶、学習、思考、予測、運動制御などの機能は脳領野の単位で実行される機能だと考えられている。また、それらの機能を数理的に記述するという方法を用いて、物理学、工学、情報学で脳科学が研究されてきた。この様に様々な分野で研究が行われており、それぞれの分野で脳の理解は異なっている。

ところで、デビッド・マー (David Marr) は、脳の情報処理を完全に理解するためには、次に挙げる3つの異なる水準についての説明が可能であり、また必要であると指摘した。

1. 計算理論 (computational theory)
2. 情報表現とアルゴリズム (representation and algorithm)
3. ハードウェアによる実現 (hardware implementation)

まず、計算理論の理解とは、脳が情報処理によって行う計算の目的、別の言葉で言えば、脳が行っている情報処理の最適性について理解することである。例え

¹学術的には「脳科学 (Brain science)」という言葉はほとんど使われておらず、神経科学 (Neuroscience) という言葉が一般的に使われる。これは主に生物学や生理学で使われてきた言葉である。他に、医学では、主にヒトの脳損傷を研究対象とした神経心理学 (Neuropsychology) があつた。認知科学では、1990年代から脳の非侵襲計測が発展してきて、脳を積極的に研究対象とするようになったが、その場合も認知神経科学 (Cognitive neuroscience) と称されており、「脳科学」は出てこない。日本では1997年に、脳の総合的な理解を目指すという理念のもと、理化学研究所 脳科学総合研究センター が設立された際に「脳科学」という言葉が使われた。本論文では、神経科学や認知科学などヒトの脳や神経を理解を目指す学問全般を指して「脳科学」を使う。

ば視覚において網膜像として写る画像から外界の特性を正確に取り出すためには、なぜその計算をするか、また、何を計算するか、を理解することである。そのためには、脳だけを調べていたのでは十分ではなく、脳にとっての外界、例えば身の回りに存在する視覚刺激の特性などを考察し、脳が解いている問題を明らかにする必要がある。また、実際の研究においては、ヒトや動物に課す認知行動の実験課題を定義することが、この水準の理解につながる。

情報表現・アルゴリズムの水準では、計算理論が実際にどの様の実現されているかを理解する。情報表現とは入出力のための脳内表現 (representation) であり、アルゴリズム (algorithm) は入出力の変換の方法である。この水準は脳科学において、特に生理学や心理学において、もっとも関心をはられ調べられてきた水準である。電気生理学では、神経活動 (主に発火率) を調べることで、各分野の神経活動の情報表現を理解しようとしているし、認知科学では刺激と行動の入出力関係から、脳で実行されているアルゴリズムについて理解しようとしている。

最後の水準であるハードウェアによる実現とは、脳の中で情報表現・アルゴリズムが物理的にどのように実現していかを理解する。例えば、物理的特性をもつ神経細胞がどの様にして情報表現やアルゴリズムを実現しているかである。また、「脳の」情報処理を理解するためには、確かに「脳で」実現していることを確認することが必須である。

このように、脳の情報処理を説明するためには、3つの異なる水準についての説明が必要であり、3つの水準の理解が得られてはじめて、脳の情報処理を理解したことになる。しかし、現在の脳科学では、先ほど特定の学問が特定の階層にある機能を研究対象としていると説明したように、各学問では解く手の水準についてを対象に研究されている現状があり、他の水準への注意が十分に払われていない。

特に、脳科学の主流である生理学や心理学では、情報表現・アルゴリズムやハードウェアについての議論は積極的に行われてきたが、計算理論の水準については十分に言及されていないように思う。脳の計算理論についての理解を目的とする、もしくは計算理論についての理解を手がかりに情報表現・アルゴリズムの研究を行う学問領域として計算論的神経科学があるが、情報表現・アルゴリズムの水準について数理的な表現でモデリングを行う学問として紹介されることも多い。計算理論の水準についての理解があることに気がつかれていない事も多いようである。数理モデリングすることで、情報表現・アルゴリズムについての理解が深まる。また、数理モデリングは計算理論の水準を考えるための大きな手がかりを得られるので、計算論的神経科学にとっても必要な作業である。しかし、それだけでは脳の情報処理を理解したことにはならない。なぜそのように数理モデリングされる情報表現・アルゴリズムが脳で行われているのかを問う、計算理論の水準

が重要であることを強調しておく。

脳を構成する組織は可塑性を持ち、その機能は外界に適応的である。普段は行わない行動についても訓練を十分に行えばできるようになることは、日常的に見ることができる。研究においても同様であり、実験参加者に実験課題の訓練を行えば、課題を遂行できるようになり、行動としてできている程度のアルゴリズム・情報表現を脳が獲得する。脳が問題を解くために採用するアルゴリズムは外界に強く影響を受ける。大脳皮質で情報処理される後天的に獲得する機能については、解かれる問題を定義せずに脳の情報表現やアルゴリズムを議論することは難しいのではないだろうか。

また、脳科学は、さまざまな学問領域で研究されており、これら多くの知見をつなげるためにも計算理論を考えることが重要であると考えている。もし、自分たちが普段行っている認知行動の情報処理について知りたければ、実環境に沿った実験課題である必要がある。しかし、実際の研究では、実環境の感覚刺激を再現することは難しく、また実環境での実験を採用したとしても多くの関連要因を切り分けて分析することが不可能であるなどの理由から、実環境を抽象化した実験課題を作成する。その場合、例えば、同じ脳の情報処理を対象にしているにもかかわらず、厳密には異なる実験課題であるため、異なる結果が得られることがある。これは、当然だが、実験課題が異なる程度に脳の情報処理も異なっていることを反映している。この問題の対処として、実験課題を統一して網羅的に実験をやり直すことで再検証するという方法がとられることもある (eg.[8])。しかし、できるならば、それぞれの研究が行った実験課題の差異を議論して、次の研究へとつなげられるようにすることが科学的な方法だと考えている。計算理論について考えることは、実験課題を分析することでもあるから、その研究の一般性と特殊性がわかり、他の研究との関連性を議論しやすくなると考えている。

計算理論の重要性として強調したことは、そのまま他の水準の重要性となることも多くあり、実のところは計算理論特有の重要性ではない。特定の水準が他の水準と比較して重要であるということはない。これら3つの水準における理解は、相互に制約を与え、ある水準について考える場合において、他の水準について理解しておくことが研究を進める上で有益となる。たとえば、計算理論が分かっているならば、その計算目的を果たすアルゴリズムや情報表現のみを考えれば良い。また、脳のハードウェアが分かっているならば、そのハードウェアで実現可能なアルゴリズムのみを考えればよい。3つの水準が互いに独立しているわけではないので、3つの説明に分けられないこともあるが、3つの水準を念頭におくことは脳の情報処理を理解するための良い指針になると考えている。階層性のある機能単位を持つ脳において、全ての階層において3つの水準の説明が行われ、階層間をつなぐ説明がなされたときに脳を理解したと言える。

本論文の構成

私は、ヒトの高次視覚情報処理について、計算理論を考慮した実験を行い、主に情報表現・アルゴリズムについての解明を目指した。

第1部の序論では、そもそも何を以て脳を理解したことになるのかを議論するために、マーの提唱した3つの水準を紹介した。特に脳科学ではまだ重要性が十分に認識されていないと感じている計算理論の水準について強調して考えを述べた。序論に続き、本論文で対象とした脳の視覚情報処理について、特に高次機能の領野についての概要を述べた。また、脳の情報処理を調べる上での問題点を明確にし、本研究で用いた方法論である行動予測について概要を説明した。

第2部では、具体的な2つの研究を挙げ、第1部で述べた目的や方法論が実際の研究でどの様に適用されたかを述べた。1つ目は眼球運動計測を用いた心理物理実験研究であり、視覚的に与えられた系列の予測について議論を行った。ダイナミクスを持ち非周期的に現れる系列を、ヒトが予測できるか、さらにどのような予測を行っているかを議論した。2つ目はfMRI脳活動計測を用いた研究であり、観測された行動と脳活動から、主に顔知覚の情報表現についての議論を行った。2つの具体的な研究を通して、行動予測をどの様に適用したか、それによって得られた知見を述べた。

第3部では、2つの研究によって得られた知見をまとめ、議論した。また、本論文で議論した方法論についての展望を述べた。

第1章 研究対象と方法論

1. 視覚系と高次情報処理

本論文では、ヒトの視覚系について議論する。大脳皮質では第1次視覚野から視覚情報処理が始まり、サルの大脳皮質では30程度の視覚領野が定義されている。ヒトにおいても、レチノトピー構造があることを利用して、fMRI研究で機能的に視覚領野が定義されている。Ungerleiderらにより、解剖学的知見と損傷研究の知見から視覚情報処理は大きく分けて2つの経路があると提唱された[66]。それ以降、様々な研究から、物体の運動など空間的・動的な視覚情報は背側経路で情報処理され、色や形など静的な視覚情報は腹側経路で情報処理されることがわかっている。ただし、背側経路の領野に色や形などの物体知覚に関する情報表現がないわけではない[34]。物体知覚に関わる情報はそれぞれの経路に存在しているが、情報処理の目的（最適性）にあわせて異なる情報表現を持つと考えるのが適切だろう。上記で述べた背側経路と腹側経路での視覚情報処理の違いとは、処理されている情報表現の差異ではなく、情報処理の最適性の差異であると考えられる。

第2部で研究対象となる眼球運動系（サッカード系）は、背側経路で処理された視覚情報をもとに、運動出力情報が生成される外側頭頂間領野（Lateral intraparietal area, LIP）や補足眼野（Supplementary eye fields, SEF）、前頭眼野（Frontal Eye field, FEF）での情報処理が行われる。第3部で対象となる顔知覚は、腹側経路の高次視覚領野にあたる紡錘状回（Fusiform Area, FA）、特に顔に選択的に活動する紡錘状回顔領域（Fusiform Face Area, FFA）に情報表現が存在するとされ、さらに顔知覚にも関わっている。本論文では背側経路と腹側経路における視覚系、特に行動（運動）や知覚・認知に近い高次視覚情報処理を研究対象とした。

本論文で議論した情報処理は、他の感覚系においても同様の情報処理が行われている可能性がある。時間系列をもつ刺激は視覚に限らずありえるから、その予測である系列予測は他の感覚系においても研究が行える。顔知覚については、個人の同定を行う場合、顔やその他の身体的特徴と性別や社会的な地位、関係性などを結びつけて行ので、視覚系だけで閉じた情報処理ではない。

それでも本論文で議論が視覚系に限定しているのは、次の研究上の利点が挙げられるからである。視覚は、他の感覚系に比較して、様々な観点で研究されてい

る。錯視などさまざまな視覚的現象が知られており、その情報処理を行っている脳に目を向けた場合に、視覚系や眼球運動系は解剖学的な結合、機能的な結合など回路の特性が、神経生理学で最も詳しく調べられてきたと言ってもいい。さらに踏み込んでアルゴリズムや情報表現の研究を行うのに最も適した感覚系であると言える。

また、眼球運動系を研究対象とすることは、腕など他の運動制御系を対象とするよりも、有利である。眼球運動系は視覚と密接な関係を持ち、視覚と同様に関係する脳の知見が多く蓄積されている。眼球運動の効果器である眼球には上直筋、下直筋、内直筋、外直筋、上斜筋、下斜筋の6つの筋があり、眼球はこの6つの筋が伸縮することで制御されている。その運動は単純であり、眼球中心を通る回転運動にほぼ限られる。特に、内外直筋が水平方向に沿った筋肉であるため、目を左右に動かす運動との対応が明確であるなど、理解しやすい制御系であるという利点もある。

2. 行動予測

脳科学の認知行動についての研究では、脳が行う情報処理の定義の多くは、実験設計として表現される作業仮説による説明である。古典的な差分解析をするための実験設計では、2つの実験課題を用意し、一方の課題では対象とする情報処理が伴う実験課題、他方を情報処理が伴わない課題にし、2つの実験課題における行動（例えば、正答率、反応時間）や脳活動の差を調べ手いる。普通は多くの作業仮説を検証し、総合的に考えて1つの情報処理の仮説が妥当だと理解される。1つの作業仮説は、対象とした情報処理そのものを表現しているわけではない。

脳科学における多くの研究者は、操作的定義や心理学で言われる傾性概念 (disposition concept)¹ について述べたいわけではなく、因果関係をも含意する記述を目指している。ただ、発表された報告が実験課題の進行や作業仮説のみの説明であると、その研究で仮説とした脳のアルゴリズムについて明にわからないことが多い。実験的検証では必然的に作業仮説を用意せざるを得ない。しかし、作業仮説は研究者がもつ計算理論と情報表現・アルゴリズムの仮説から演繹的に推論される。脳の情報表現やアルゴリズムの仮説を明に示した上で、そこから導かれた作業仮説を説明することが他の研究へとつなげるために重要である。

物理学では自然現象を記述するために数式を用い、定量的な評価や複雑な推論を行うことで発展してきた。工学や情報学でも情報処理機械を構築するために数

¹観察された行動を記述しただけの概念であり、概念の内容は観察に完全に還元される。操作的定義のされる概念の多くは傾性概念。一方で、観察に還元できない剰余意味 (surplus meanings) を持っており、行動の因果的な法則性なども含む概念は理論的構成概念 (Theoretical construct) もしくは仮説構成概念 (Hypothetical construct) である。

理的な記述が用いられてきた。計算過程であるアルゴリズムや情報表現を記述するためには、自然言語によりも、数理的記述が適していることが経験的に示されてきた。脳の情報処理を解明する研究においても、アルゴリズムと情報表現にを数理的に記述することが研究を発展させる上で重要だと考えている。

そして、数理的に記述されたモデルを実験的に検証するための手続きとして予測がある。赤池が提唱した情報量基準 (An Information Criterion; AIC²) など、仮説とした数理モデルを定量的に評価する方法が数理科学で様々に考えられている。機械学習の分野では、性能評価は交差検証法 (交差確認法; cross validation) などによる予測性能評価 (汎化性能評価) を示すことが多い。科学においては、「現象を良く予測すること」自体は目的ではないが、予測できる程度に現象を記述し、予測することで理解を確認する事は重要な手続きであると考えた。

認知行動の研究では、実験参加者に視覚刺激など感覚入力を与え、行動 (運動) 出力を観測し、感覚入力-行動 (運動) 出力の変換を考えることで、ブラックボックスである脳の情報処理を解明しようと試みてきた。ここで数理モデルによって予測すべきは行動であり、予測した行動と観測された行動を比較することで、数理モデルを評価することができる。

視覚系の研究では、脳活動から視覚運動刺激の識別行動を定量的に予測した最初の研究として、Newsome らが中心となって行った一連の研究が挙げられる [42, 5, 52]。

本研究で議論する2つの研究では、実験課題の計算理論を明らかにし、もしくは、ヒトの脳で行われている情報表現・アルゴリズムの仮説を明に定義した。それらの記述は数理的術を用いた。1つ目の研究での実験課題で、実験参加者は視標系列を見て、次時刻の視標位置を予測した。2つ目の研究では、実験参加者は1つの顔を見て、次に表示される2つの顔のどちらに似ているかを判断した。それらの実験参加者の行動を予測することにより、検証を行った。

²赤池情報量基準 (Akaike Information Criteria) とも呼ばれるが、この情報量基準を提唱した赤池自身は一貫して AIC の A は冠詞の A だと主張しており、本論文ではそれに習った。

第2部

高次視覚情報処理の研究

眼球運動計測を用いた系列予測の研究

第2章 系列予測の研究背景と目的

1. 予測のための脳情報処理

私たちの脳には生まれたときから備わっている固定的な機能もあるが、それ以上に与えられた環境を学習して適応する機能が多いと考えられる。

環境に適応的な機能の一つに予測がある。手を伸ばして目標を指し示す到達運動課題において、系列運動制御や系列運動学習などの研究が行われている [25, 15, 64, 9, 48]。そのなかでも小川らは視覚フィードバック遅延を入れた視覚誘導運動の fMRI 研究で運動予測を行っている脳領域を示している。

眼球運動はシステムが単純であり、視覚運動制御の研究に理想的な系であるため、予測に関する研究にしばしば用いられてきた [3, 63, 33, 15, 41]。運動物体を見るために、眼球運動によって視覚的に追跡するという行動がある。私たちが住む環境には物体が様々な動きを見せるが、その動きを眼球運動によって追跡するためには、その制御に予測が不可欠であると考えられる。

例えば、円滑性追跡眼球運動 (Smooth Pursuit Eye Movements; SPEMs) は速度 10 deg/s 程度の滑らかにゆっくり運動するターゲットを遅延無しで精度良く追跡することが可能であることが知られているため [68, 12, 61]。一方でヒトは、物体を正確に視認できる中心視が視野角 2 deg 程度の狭い範囲に限られており、さらに視覚入力から眼球運動出力されるまでに 250 ms 程度の遅延がある。先ほど述べた視標速度 10 deg/s は、遅延時間 250 ms の間に視標が視野角 2 deg から外れてしまう速度であり、この遅延時間を補償する機構が不可欠であり、それが予測だと考えられる [51, 68, 12, 46]。また、理論的には何らかの予測制御をしていると考えられている [2, 58]。

上述のように、予測の情報処理がさまざまな運動系に存在している。本研究で対象でもあるサッカード (saccade) と呼ばれる眼球運動系においては、ハードウェア上の制約などがなく、予測が必ずしも必要でない場合にも予測的な眼球運動が起こることが知られている [57, 31, 20]。例えば、画面上の 6 つの異なる位置にある点が決まった順に点灯し、被験者は視標が点灯したら、その位置に眼球運動するという課題において、試行を繰り返すことで点灯する位置と順序が学習されると、点灯する前に予測的な眼球運動を行うと報告された [33]。

ここで、サッカードが駆動する際に関わる脳の領野とその情報表現は、電気生

理学研究から、上丘 (Superior Colliculus, SC) が関与している事が知られている。特に SC はサッカード生成に重要な役割を果たしており [39], 眼球の空間位置と運動の大きさの情報表現があることが知られている [35, 21, 45] また、サッカードに先立ったニューロン発火が起きていることも知られている [22, 40, 13, 14]. 大脳皮質では、前頭眼野 (Frontal Eye field, FEF) [23, 50, 6, 53, 54] 補足眼野 (Supplementary eye fields, SEF) [55][53], 頭頂間溝外側部 (Lateral intraparietal area, LIP) の 3 つが領野が関連している。前頭眼野と頭頂間溝外側部は上丘への直接の投射があることが知られており [27], 前頭眼野が、電気刺激でサッカードが誘発されたり [4], 損傷によってサッカードの振幅が減少する事が知られている [19]. 一方で、補足眼野は直接には運動出力にはつながっていない領野であり、学習などによって神経活動が変化する領域である事が知られている [11, 10]. また、補足眼野を損傷した患者を対象とした研究で、運動系列の学習や実行が阻害されたという報告もある [47]. ただし、サッカード系における予測がどの領野で行われているかを示した研究はまだない。

サッカード系において、次時刻の位置を予測する行動は、記憶した過去の位置や大きさの情報から予測している運動しているだろうと考えられる。さらに系列を予測する行動は、位置や大きさに時間順序を加えた情報、つまり時間的パターンの情報を用いて予測しているだろうと考えられる。

ところで、視覚眼球運動系の先行研究で用いられてきた視標系列は、動きの要素である位置や時間的パターンの周期性を記憶して、それを単に繰り返すことで最適な予測が可能な系列であった。しかし、ヒトが住む実環境では、力学的な相互作用の結果、複雑な時間発展のある法則性 (以後、ダイナミクスと呼ぶ) をもつ動きが現れる。ヒトはダイナミクスをもつ運動を目で追うという行動を日常的に行っているように思える。またそうだとした場合に、計算論的に考えて、物体運動がもつダイナミクスを学習して、そのダイナミクスをもって予測することがうまく予測方法である。しかし、サッカード系において系列のダイナミクスの学習や予測という観点で調べられた研究はなかった。

2. 本研究の目的

本研究では、単に位置や周期性の情報を用いた予測ではなく、位置や大きさのダイナミクスを学習した上での予測をヒトが行うことを、眼球運動計測を用いた心理物理実験から調べた。また、ヒトの予測系列を解析することで、どの程度の過去の情報 (視標位置) を用いて予測しているかを調べた。

位置やその周期性のみでは表現されない、ダイナミクスをもった視標の動きをヒトが予測できるかを検証した。さらに、ヒトが、どの程度の過去にまでさかの

ぼって視標位置を参照して予測しているのかなど，ヒトの予測行動の特性を検証した．

第3章 実験方法

1. 系列予測サッカード課題

本研究では、実験参加者は提示された現時刻の視標を注視した後、次時刻の視標位置へ能動的にサッカードさせる系列予測サッカード課題 (Sequential predictive saccade task) を用いて、ヒトの系列予測について調べた。

まず、実験参加者が行ったタスクの最小単位である 1 試行の定義を述べる (図 3.1)。第 $T+1$ 試行において、モニタ上に視標 $S(T)$ が表示される。実験参加者は表示された視標 $S(T)$ を注視する (Target gaze, 視標注視)。視標は 1500 ms 間モニタ上に表示されたあと、合図の音とともに消失する。実験参加者は視標 $S(T)$ の消失後ただちに、次試行で現れる視標位置 $S(T+1)$ を予測し、予測位置 $P(T+1)$ を注視する (Prediction gaze, 予測注視)。視標の消失時間は 1500 ms である。この視標注視と予測注視を 1 度ずつ行うことで 1 試行とした。予測注視終了時に合図として再び音が鳴る。1000 ms をおいて次試行の視標位置 $S(T+1)$ が表示され、次試行が開始される。視標位置 $S(T+1)$ と予測位置 $P(T+1)$ が異なっていれば、実験参加者には予測誤差が知覚される。この試行を繰り返すことで、実験を行った。実験参加者へは、後述する系列の特性については教示せず、次試行の視標位置を予測することのみを教示した。

また、視標位置は、視野角 (deg) で表現し、本実験では視標が表示される範囲は ± 5 deg 以内になるように設計した。これは眼球運動の線形性が仮定できる程度に小さい視野角の範囲である。

2. 実験手順

1 つの系列について実験である 1 タスクは全 500 試行で構成される (3.2)。実験参加者は 1 ブロック分の 25 試行を連続して行い、ブロック間に適宜休憩を挟み、20 ブロックを行い、1 タスク 500 試行を終了した。

実験参加者は、系列予測サッカード課題を繰り返し 25 試行を行い、これを 1 ブロックとした。1 ブロックは 100 s 程度で終了するが、数ブロックを続けて行

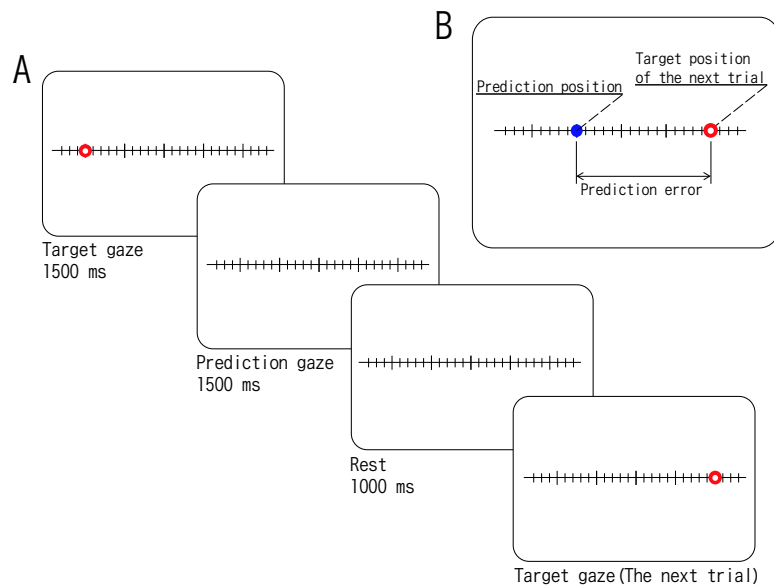


図 3.1: 系列予測サッカード課題：1 試行.

図 A：実験課題 1 試行でモニタ上に表示される視覚刺激である．まず，モニタ上に視標 $S(T)$ が表示される．実験参加者は表示された視標 $S(T)$ を注視する（Target gaze, 視標注視）．1500 ms の後，合図の音とともに視標 $S(T)$ が消失する．実験参加者は視標 $S(T)$ の消失後ただちに，次試行で現れる視標位置 $S(T+1)$ を予測し，予測位置 $P(T+1)$ を注視し，その位置から目を動かさないようにする（Prediction gaze, 予測注視）．予測注視時間は 1500 ms である．この視標注視と予測注視を 1 度ずつ行うことで 1 試行とした．

予測注視終了時に合図として再び音が鳴る．1000 ms をおいて次試行の視標位置 $S(T+1)$ が表示され，次試行が開始される．この時，実験参加者は 1 試行前の予測位置 $P(T+1)$ から表示された視標位置 $S(T+1)$ へ目を動かすことになり，予測誤差が知覚される．

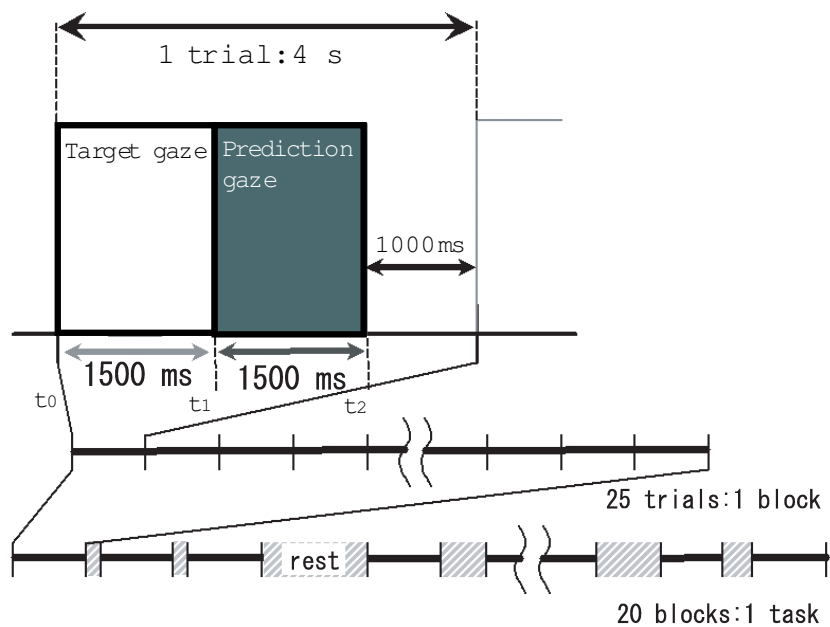


図 3.2: 系列予測サッカード課題・実験進行.

実験を構成する 1 試行, 1 ブロック, 1 タスクの進行を図示した. 1 試行は 4 sec であり, 連続して 25 試行を行い, 1 ブロックが終了する. 1 ブロックごとに数十秒から数分の休憩を挟み, 20 ブロックを行い, 1 タスク全 500 試行を終了した.

予測注視 (Prediction gaze) 後に 1000 ms の時間が設定されているのは, この時間に実験参加者がまばたきを行えるようにするためである.

うと瞬きが多くなり、正確な計測・分析が困難となるため適宜ブロック間に休息を入れた。1つのタスクは適宜休憩を挟みながら最大3時間ほどで終了した。

3. 実験参加者

実験に、KT, MO, NN, TY, KY の5名の20代男性に実験参加者として参加してもらった。AR₁タスクと対をなすランダムタスクにKT, NN, MO, TYの4名、AR₂₁, AR₂₂と対をなすランダムタスクにKY, NN, MOの3名を実験参加者として行った。実験参加者NN, MOが全てのタスクに参加した。3つの実験全てに参加した実験参加者は3名であり、2つの実験に参加した実験参加者は1名であった。全実験参加者は正常な視力であるか、矯正によって十分な視力を持つ実験参加者であった。また、全実験参加者は実験課題について、教示した内容以外は知らされていない。

4. 実験設定の詳細

4.1 視標提示

視標は半径0.16 degの赤円の中心に半径0.06 degの黒円で表される中空円として描画した。実験参加者には視標の中心にある黒円を注視するように教示した。また、ディスプレイに表示された黒の背景の上に、視標が現れる高さに水平線を表示し、さらに水平線上±5 degの範囲で0.36 deg刻みに垂直線を描画し、格子点を表示した。目盛り線は実験中に常に表示した。実験中に実験参加者が何も表示されない背景画面を繰り返し注視する場合では、非常に疲労し実験課題の遂行が困難になることがあるため、目盛り線を提示して注視位置を明示することで改善した。

後述する視標時系列の作成においては、まず式(4.1)もしくは式(4.2)に従って時系列データを作成し、その後、その時系列の各値が最も近い目盛り線上にくるように値を量子化し、これにより視標時系列の値を27個に離散化し、視標は水平線と垂直線の交点である目盛り格子上のみに現れるようにした。

4.2 実験機器と解析装置

眼球運動計測にはEyelink II (SR Research 社製)を用い、サンプリング周波数には250Hzもしくは500Hzで計測した。視標は15.3 inchのTFTカラー液晶モニタ(解像度、1280x1024 pixels; リフレッシュレート、60 Hz)に表示し、モ

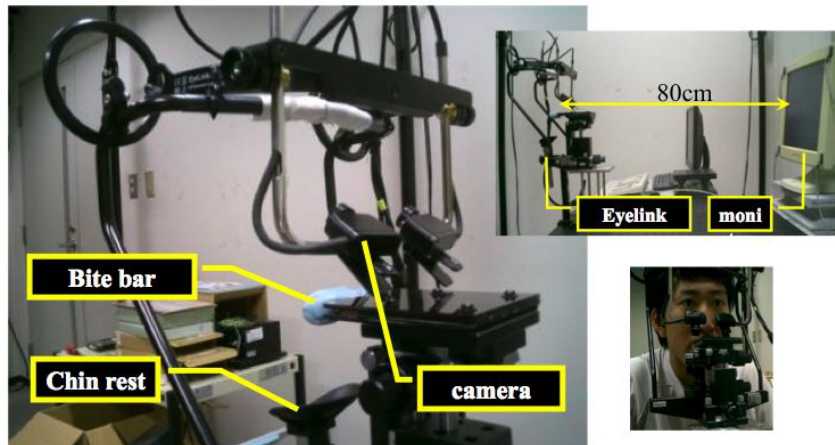


図 3.3: 眼球運動計測装置. 実験における視線位置計測には眼球運動計測装置 Eyelink II を用いた (左図). camera 部分で眼球運動を撮像し, モニタ上の実験参加者が見ている位置を計算する. 実験参加者の頭部は Bite bar と Chin rest によって固定した. 実験参加者の眼球からモニタまでの距離は 80 cm である (右上図).

ニタの応答速度 (暗→明→暗) は 80 ms である. 視標表示は Dell 社製, Pentium III-746Hz によって制御した. 実験参加者の前額面からモニターまでの距離は 80 cm の距離とし, 実験参加者の頭部は動かないように顎台およびバイトバーによって固定した. 取得した視線位置のデータは Matlab (Mathworks 社) を使用した.

4.3 計測と前処理

計測において予測位置は視標消失後 500 ms の時の視線位置とした. これは 300 – 400 ms の反応時間と, 十分に長くとしたサッカードの持続時間 100 ms を仮定したからである. つまり, 眼球運動系の議論が不要である. その上で, 実験参加者に 1 試行毎の予測を無意識的に行わせることのできる時間設定であると考えた. また, 実験参加者には 1 試行において 1 度予測して視線を動かした後に再度視線を動かさぬように教示した. また, 実験参加者には, 目盛り格子上のみに視標が現れるので予測注視時には目盛り格子上を注視するようにと教示している.

前述のように, 実験参加者には, できる限り予測注視後の 1000 ms の空白時間に瞬きをするように教示しているが, 教示した以外の時間帯にも瞬きが観察された. 特に, 視標注視予測注視の両注視前後に起こる眼球運動開始・終了時に特に瞬きが多く観察された. そのため視標注視時間, 予測注視時間の 1500 ms の前後 500 ms を削除し, 504 ms-1000 ms の計測値を抜き出して解析した. また, 抜き出した 500 ms-1000 ms 間においても瞬き時のデータが入る可能性を考慮し, 抜き出した計測の中央値をとり安定的に注視座標を取得するようにした (計測値抽出).

計測値を較正するために試行毎に行う視標注視を利用する。実験参加者は試行毎に視標注視を行っている。この時の視標表示した視標座標 $S(T)$ と視標注視時間に計測された注視点座標 $E(T)$ との差 $S(T) - E(T)$ をとる。予測注視時間に計測された注視点座標 $P(T)$ に $S(T) - E(T)$ を加え予測注視座標 $P'(T) = P(T) + S(T) - E(T)$ とする（計測値補正）。以上の計測値抽出と計測値補正を解析の前処理として行った。本研究の実験で用いた視標と目盛り線を使用した計測精度を確かめる予備実験において、本実験の計測誤差は 0.11 deg であることがわかっている。

第4章 視標系列の計算理論

本章では、予測の計算理論を明らかにするために、実験参加者が予測した系列の最適性など、系列の特性について述べる。まず、ダイナミクスを持つ系列として実験課題に採用した自己回帰過程 (Auto-Regressive process. 以下, AR 過程) を説明し、あわせて対照実験とした乱数系列の作成について説明する。次に, AR 過程について解析的な最適解など系列の特性を述べる。さらに, 実際に実験で用いた系列は, 用いていた乱数が疑似乱数であること, 時系列が500 試行分と短いこと, 時系列を離散化していることなどの要因により, 解析的に調べた結果と厳密には一致しないので, 数値計算 (計算機実験, シミュレーション) を行った結果を示し, 系列の特性について詳細に議論する。

1. 視標系列

1.1 自己回帰過程

ダイナミクスを持つ視標時系列の生成モデルには, AR 過程を採用した。本研究では, 式 (4.1) で表される1次 AR 過程の系列を用いた1つの実験 (AR_1 タスク) と式 (4.2) で表される2次 AR 過程において異なる2組のパラメータセットを用いた2つの実験 (AR_{21} タスク, AR_{22} タスク) の計3つの系列についての AR タスクを行った (4.1)。

$$S(T) = a_1 S(T-1) + \xi(T). \quad (4.1)$$

$$S(T) = a_1 S(T-1) + a_2 S(T-2) + \xi(T). \quad (4.2)$$

ここで, システムノイズ項 $\xi(T)$ は平均 $\mu_E = 0$ deg, 標準偏差 $\sigma_E = 0.72$ deg のガウス分布に従うとした。 a_n は n 次項の AR パラメータであり, 本研究においては具体的に AR_1 タスクで $a_1 = 0.9$, AR_{21} タスクで $a_1 = 0.6, a_2 = 0.3$, AR_{22} タスクで $a_1 = 1.3, a_2 = -0.7$ と設定した。

$$S(T) = 0.9 S(T-1) + \xi(T). \quad (4.3)$$

$$S(T) = 0.6 S(T-1) + 0.3 S(T-2) + \xi(T). S(T) = 1.3 S(T-1) - 0.7 S(T-2) + \xi(T) \quad (4.4)$$

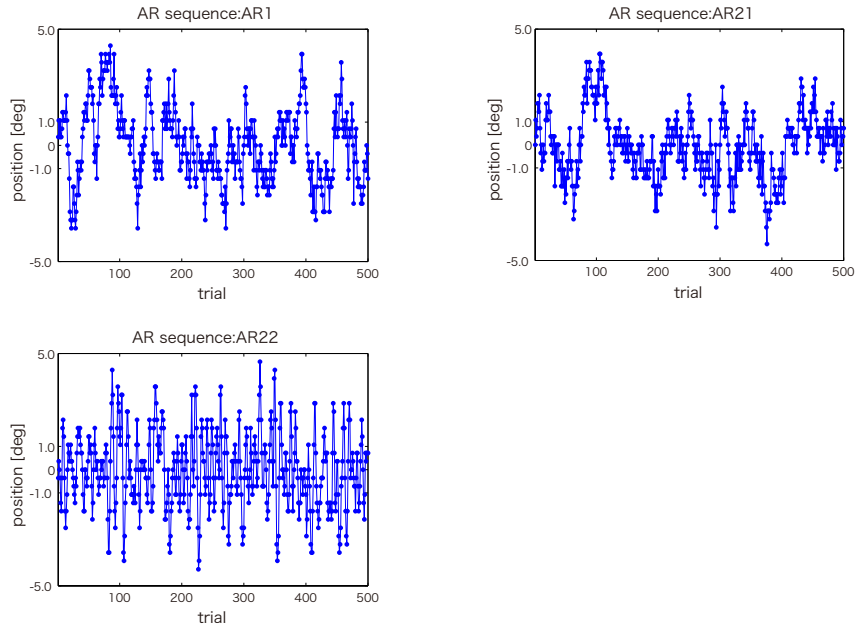


図 4.1: AR 過程の系列. AR_1 タスクで用いた 1 次 AR 過程の系列 (左上), AR_{21} タスクで用いた 2 次 AR 過程の系列 (右上), AR_{22} タスクで用いた 2 次 AR 過程の系列. 横軸: 試行数, 縦軸: 視標位置 (deg).

AR 過程は非周期的であり, 周期的なパターンを単純に記憶する方法では一般に予測不可能である. 一方で, 予測誤差が最小となる最適予測時の予測方法と予測誤差は解析的に明らかにされており, 実験結果を定量的に解析することが可能となる.

1.2 乱数系列

対照実験となる乱数タスクの時系列は, 各タスク毎に用いた AR 過程の時間順序をシャッフルすることによって生成した. これにより, 乱数時系列の平均, 標準偏差 $SD_{rnd}\{S(T)\}$ など統計量は AR 過程と等しいが, 時間順序はランダムであるため予測が不可能な系列を作成した. 乱数タスクは, 3 つの AR タスクに対してそれぞれ生成しており, AR_1 タスク, AR_{21} タスク, AR_{22} タスクの対となる乱数タスクをそれぞれ, rnd_1 タスク, rnd_{21} タスク, rnd_{22} タスクとした.

2. 視標系列の分析

2.1 予測誤差：最適予測時の予測誤差

T 試行時における予測を $P(T)$ ，最適予測を $\hat{P}(T)$ ，予測誤差を $E(T) = \hat{P}(T) - P(T)$ とし，予測誤差の標準偏差を

$$SD\{E(T)\} = \sqrt{\frac{1}{N-1} \sum_{n=1}^N (E(n) - \mu)^2} \quad (4.5)$$

とする． $\mu = 1/N \sum_{n=1}^N E(n)$ である． N は各タスクの試行回数であり，本研究では 500 である．AR タスクと乱数タスクを区別するために，それぞれの予測誤差の標準偏差を $SD_{ar}\{E(T)\}$ ， $SD_{rnd}\{E(T)\}$ と表記する．

最適予測 $\hat{P}(T)$ は，1 次 AR タスクでは $\hat{P}(T) = a_1 S(T-1)$ ，2 次 AR タスクでは $\hat{P}(T) = a_1 S(T-1) + a_2 S(T-2)$ であることが解析的に求められる．これは，実験課題において，実験参加者が視標時系列の生成モデルの線形ダイナミクスを用いて予測している場合であり，最適予測時の予測誤差 $E(T) = \hat{P}(T) - S(T)$ は生成時に用いたシステムノイズと一致し，予測誤差の標準偏差 $SD_{ar}\{E(T)\} = \sigma_E$ となる．

理論的には，被験者が予測できていれば $SD_{ar}\{E(T)\}$ が σ_E と $SD_{rnd}\{S(T)\} \doteq 2\sigma_E$ の間の値をとり，予測できていなければ $SD_{rnd}\{S(T)\} \doteq 2\sigma_E$ 以上の値をとる．

次に，AR タスクの視標時系列の生成モデルは式 4.1，4.2 であるが，システムノイズの生成に用いている乱数が疑似乱数であること，時系列が 500 試行分と短いこと，時系列を離散化していること，から作成された時系列は厳密に生成モデルと一致しない．そこで，作成した AR 系列について，システムノイズの標準偏差，視標時系列の標準偏差，AR 次数，AR パラメータを推定し，実際に用いた系列の計算理論を定義した．

まず，システムノイズと AR 系列の標準偏差を最小二乗法により求めると，システムノイズの標準偏差 σ_E は， AR_1 タスクで 0.74 deg， AR_{21} タスクで 0.74 deg， AR_{22} タスクで 0.76 deg であった．AR 過程の標準偏差 $SD_{ar}\{S(T)\}$ は， AR_1 タスクで 1.61 deg， AR_{21} タスクで 1.41 deg， AR_{22} タスクで 1.63 deg であった (4.1)．つまり，今回の実験では，最適予測である場合の予測誤差は，ここで示したシステムノイズの標準偏差の値となる．

2.2 AR 次数

同様に，理論値と一致しない可能性があるので，生成した系列について情報量規準を基にした AR 次数推定を行った．視標系列について AR 次数を求めた．AR 次

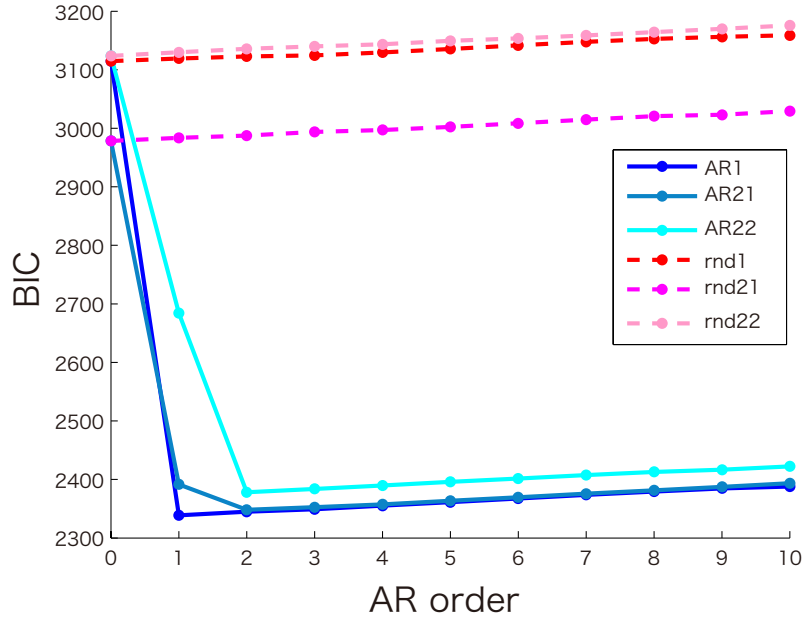


図 4.2: 視標系列の BIC 値.

本研究で用いた AR 系列と乱数系列について、各次数の AR モデルにおける BIC の値を求めた。BIC の最小値は、 AR_1 系列では 1 次、 AR_{21} 系列では 2 次、 AR_{22} 系列では 2 次と、生成モデルで設定した次数になっていることがわかった。また、乱数系列についても全て、生成モデルで設定した 0 次が BIC の最小値となった。また、3 つの乱数系列は対となる AR 系列から作成しているので、対応する AR 系列と乱数系列どうしでは 0 次の BIC 値は等しくなった。

数同定には、ベイズ情報量規準 (Bayesian Information Criteria, BIC) $BIC(p) = N \cdot \log \sigma_p^2 + p \cdot \log N$ を用いた [1, 56]。各タスクについて 0-10 次の AR モデルで、BIC を算出し、最小となる次数を予測時系列の AR 次数とした。ここで p は予測時系列を同定した時の AR 次数、 σ_p^2 は同定された AR 過程のもつノイズ項の分散である。その結果、 AR_1 タスクで 1 次、 AR_{21} タスクで 2 次、 AR_{22} で 2 次であった (表 4.1, 図 4.2)。乱数系列についても同様の解析を行ったところ、実験に用いた全ての系列で 0 次、つまりノイズ項のみのモデルであり、ダイナミクスを持たなかった。

2.3 AR パラメタの推定

AR パラメータについて最小二乗法により算出した。AR 過程のパラメータは、 AR_1 タスクで $a_1 = 0.89$ 、 AR_{21} タスクで $a_1 = 0.58, a_2 = 0.31$ 、 AR_{22} タスクで $a_1 = 1.29, a_2 = -0.68$ であった (表 4.1)。また、AR タスクと対になるそれぞれの乱数タスクについては AR 次数 0 と推定され、1 次以上のダイナミクスを持た

ないことを確認した。

2.3.1 AR パラメータの設定について

1 次 AR 過程が次の式で与えられたとき、

$$S(T) = aS(T-1) + \xi(T), \quad (4.6)$$

1 次 AR 過程の系列がもつ分散 $Var\{S(T)\}$ は、

$$Var\{S(T)\} = Var\{aS(t-1) + \xi(t)\} \quad (4.7)$$

$$= Var\{a\{aS(t-2) + \xi(t-1)\} + \xi(t)\} \quad (4.8)$$

$$= Var\{a^2S(t-2) + \{a\xi(t-1) + \xi(t)\}\} \quad (4.9)$$

$$= \sum_{i=1}^t a^{2(i-1)} \cdot Var\left\{\sum_{i=1}^t \xi(t-(i-1))\right\} \quad (4.10)$$

$$= \sigma_E^2 \sum_{i=1}^t a^{2(i-1)} \quad (4.11)$$

$$= \sigma_E^2 \cdot \frac{1 - 2^{2(t-1)}a^2}{1 - a^2} \quad (4.12)$$

$$= \frac{\sigma_E^2(1 - a^{2t})}{1 - a^2} \quad (4.13)$$

$$\xrightarrow{t \rightarrow \infty} \frac{\sigma_E^2}{1 - a^2} \quad (4.14)$$

$$(4.15)$$

となり、1 次 AR 過程の系列がもつ標準偏差 $SD_{ar}\{S(T)\}$ は、

$$SD_{ar}\{S(T)\} = \sqrt{\frac{\sigma_E^2}{1 - a^2}} \quad (4.16)$$

となる。同様に、高次 AR 過程の標準偏差 $SD_{ar}\{S(T)\}$ は

$$SD_{ar}\{S(T)\} = \sqrt{\frac{\sigma_E^2}{1 - (\sum_{i=1}^n a_i)^2}} \quad (4.17)$$

となる。つまり、AR 過程の系列がもつ標準偏差 $SD_{ar}\{S(T)\}$ は、システムノイズの持つ標準偏差 σ_E と AR パラメータによって決まる。

最適予測時の予測誤差である σ_E が視標系列 $SD_{ar}\{S(T)\}$ と同程度である場合には、最適な予測をしたとしても、実験における計測誤差などによって有意な差が認められないことがある。そのため、上記で述べた AR パラメータは、以上の解析計算から、 $SD_{ar}\{S(T)\}$ が σ_E の 2 倍程度、差にして 0.72 deg 程度となるように設定した結果である。

表 4.1: 実験で用いた系列の特性. 3つのタスクで用いた生成系列の特性を示した. AR モデルにおけるノイズ項の標準偏差である σ_E はすべての系列で, 理論的に設定した 0.72 に近い値となっていることがわかった. また, その他の特性も, 理論的に設定した数値にほぼ等しいことが確認された.

	AR_1	AR_{21}	AR_{22}
σ_E [deg]	0.75	0.74	0.76
$SD_{ar}\{S(T)\}$ [deg]	1.61	1.41	1.63
AR order	1	2	2
AR parameter	$a_1 : 0.89$	$a_1 : 0.58, a_2 : 0.31$	$a_1 : 1.29, a_2 : -0.68$

2.4 最適予測の場合の行動

逐次最小二乗法によって試行を繰り返す事により AR パラメータがどのように推移するかを観察した, ここでは, 3つの系列の比較を簡単にするため, AR_1 系列も, 他の系列と同様に 2 次 AR モデルをもちいてパラメータの推定を行った.

忘却係数 $\lambda = 1$ の場合, 3つの AR 系列ともに 100 試行程度で設定した AR パラメータに収束していることがわかった (図 4.3). これは, ある試行数における AR パラメータは, その試行数までのデータ全てを用いて最小二乗法を行った場合に推定される AR パラメータに等し. つまり, 過去の系列をすべて記憶して最適な予測を行った場合に推定された AR パラメータである. また, 予測誤差の推移を見ると, 一定で推移した. 今回の計算機実験では, 初期値を生成モデルの AR パラメータから非常に大きくはずれた値を用いた場合には, 最初の数十試行では大きな予測誤差を取る場合があるのみであった. AR パラメータの学習が予測誤差に反映されることを期待したが, そのような現象は起きなかった.

また, 乱数系列の場合にも, 100 試行程度で AR パラメータが 0 に収束していることが観測された (図 4.4)

忘却係数 $\lambda = 0.97$ の場合, 忘却係数 $\lambda = 1$ の場合ほど推定値は安定しないが, 生成モデルの AR パラメータに収束していく様子がわかった (図 4.5). 予測誤差の推移をみると, 忘却係数 $\lambda = 1$ の場合と同様に, ほぼ一定の値で推移しており, AR パラメータの学習は予測誤差には反映されなかった.

2.4.1 正弦波関数モデルを持ちいた AR 系列の予測

周期性をもつ代表的なモデルである正弦波関数をもちいて, AR 系列を予測させ, 予測誤差を調べた. 予測誤差は, AR モデルをもちいて最適予測した場合のほぼ倍になり, 一定周期をもつ予測モデルでは, 最適な予測はできないことがわかった.

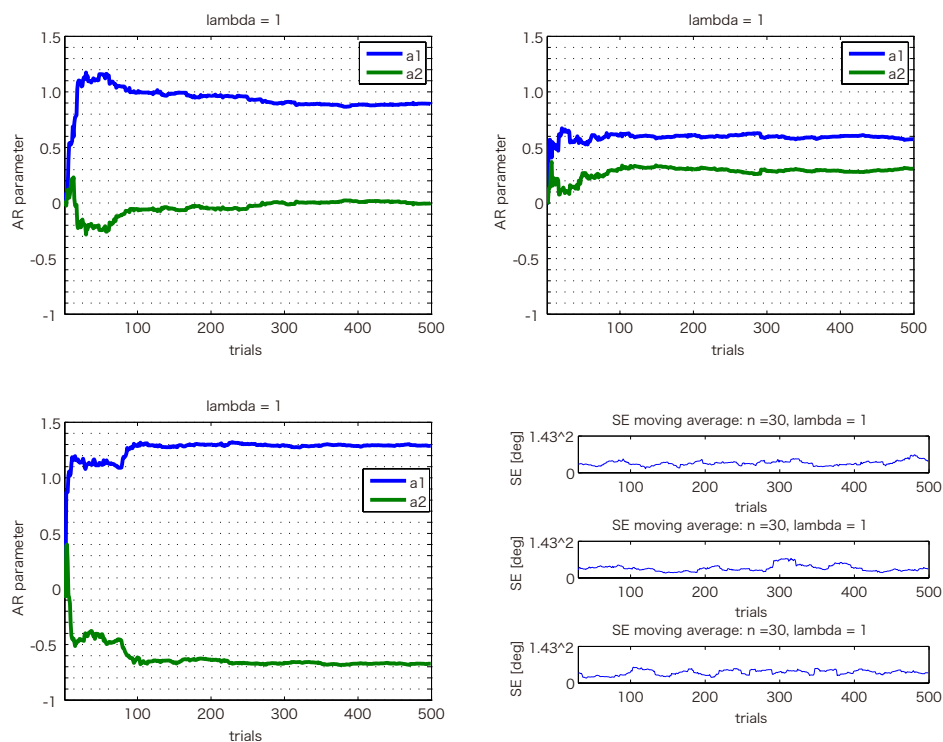


図 4.3: 逐次最小二乗法による計算機実験：AR 系列（忘却係数：1.00）。

実験課題に用いた 3 つの AR 系列について、2 次 AR モデルで逐次最小二乗法（忘却係数 $\lambda = 1.00$ ）による計算機実験を行い、推定された AR パラメータと誤差を求めた（左上図， AR_1 系列；右上図， AR_{21} 系列；左下図， AR_{22} 系列；右下図 3 つ，誤差，上から AR_1 系列， AR_{21} 系列， AR_{22} 系列）。青線が 1 次の AR パラメータであり，緑線が 2 次の AR パラメータである。いずれの系列でも AR パラメータは 100 試行にいたる前に収束していることがわかる。また，誤差については，ほぼ一定で推移した。AR パラメータの初期値を生成モデルの値から大きく変化させた場合以外は，AR パラメータで観測された学習の様子を誤差で観測することはできなかった。

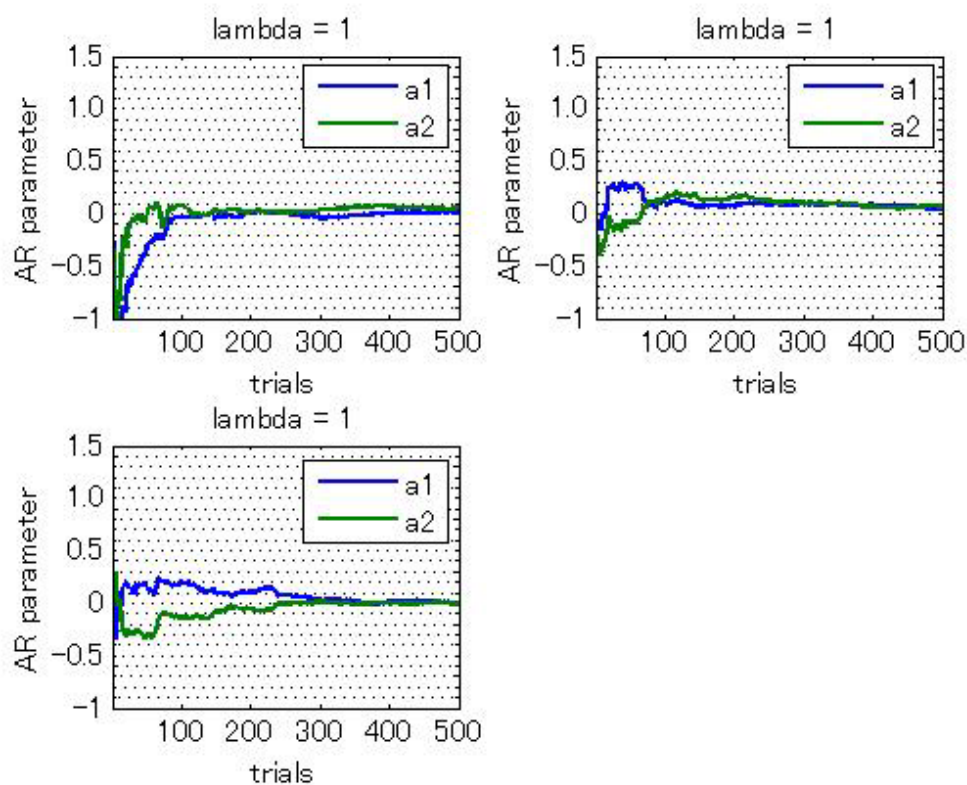


図 4.4: 逐次最小二乗法による計算機実験・乱数系列

実験課題に用いた 3 つの乱数系列について、4.34.5 と同様に、2 次 AR モデルで逐次最小二乗法 (忘却係数 $\lambda = 1.00$) による計算機実験を行い、推定された AR パラメータを求めた (左上図, rnd_1 系列; 右上図, rnd_{21} 系列; 左下図, rnd_{22} 系列)。青線が 1 次の AR パラメータであり、緑線が 2 次の AR パラメータである。2 つの AR パラメータ共に 0 に向かって収束することがわかった。

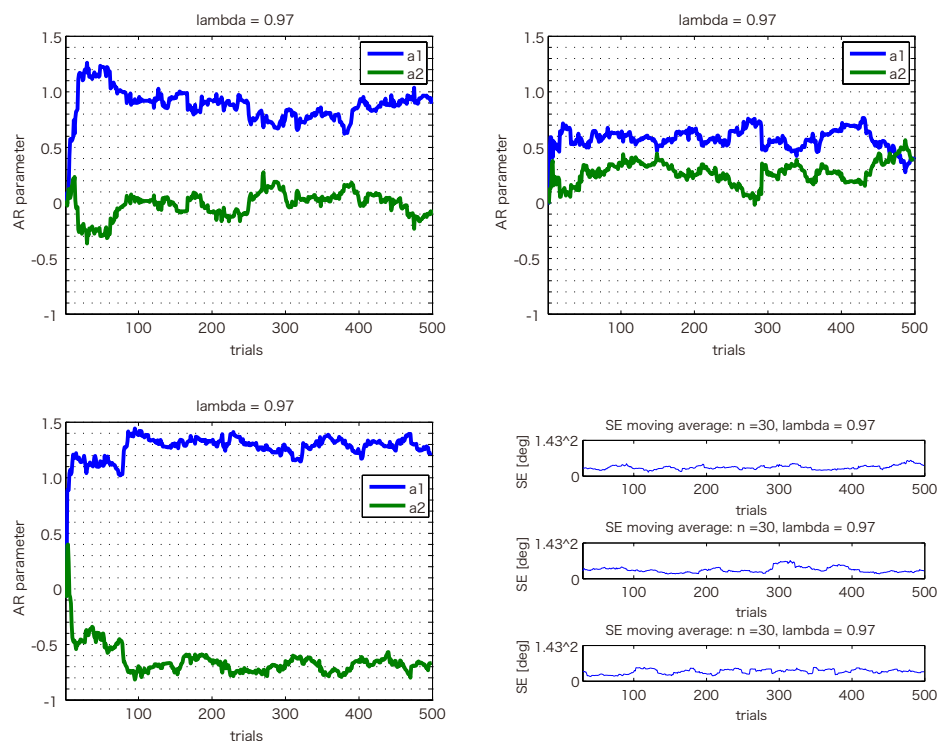


図 4.5: 逐次最小二乗法による計算機実験・AR 系列 (忘却係数: 0.97) .

忘却係数 $\lambda = 0.97$ の条件で, 4.3 と同様に, AR 系列について 2 次 AR モデルで逐次最小二乗法による計算機実験を行った (左上図, AR_1 系列; 右上図, AR_{21} 系列; 左下図, AR_{22} 系列; 右下図 3 つ, 誤差, 上から AR_1 系列, AR_{21} 系列, AR_{22} 系列). 青線が 1 次の AR パラメータであり, 緑線が 2 次の AR パラメータである. 忘却係数 $\lambda = 0.97$ の条件では, 40 試行前にかかる影響が 0.3 以下となるため, 推定値は安定しない. その他, 定性的には 4.3 の結果と同様であった.

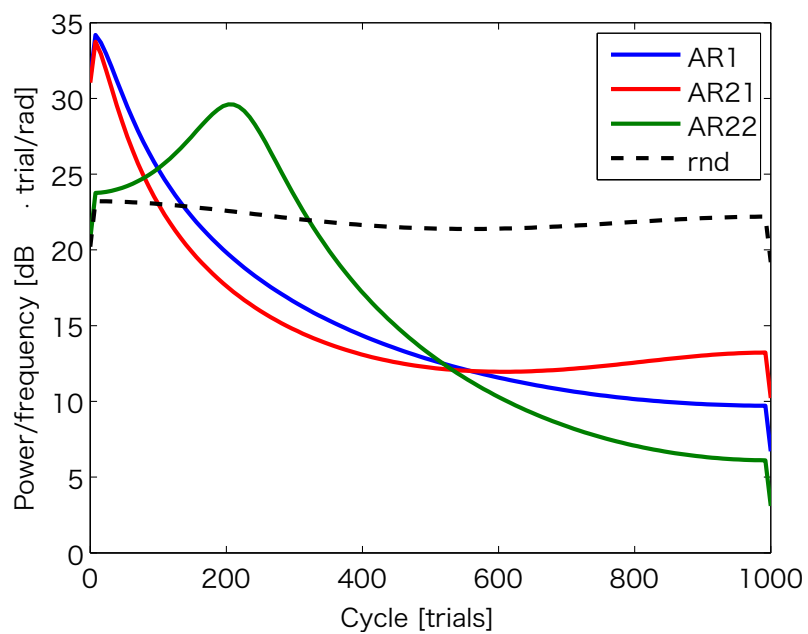


図 4.6: 視標系列の周波数分析. 本研究で用いた系列について, 周期性を調べた. AR_1 系列 (青線), AR_{21} 系列 (赤線) は最低周波数にピークを持つ. ただし, AR_{21} 系列の方が高周波領域におけるパワーが大きい. 一方で, AR_{22} 系列 (緑線) は前者 2 つの系列よりは高い周波数領域にピークを持つ. 乱数系列 (黒点線) は, 全周波数領域にほぼ一様のパワーを持った. ここで乱数系列は具体的には rnd_1 である.

正弦波モデルを予測モデルとした場合に AR 系列を予測できないことは、AR 系列をパワースペクトル解析からもわかる。AR 系列は、ピークをもつ周波数領域以外にも連続的にある程度のパワーを持つので、スペクトルピークを1つしかもたない正弦波モデルでは最適な予測はできない（図 4.6）。各 AR 系列についてパワースペクトルの特徴を述べると、 AR_1 系列と AR_{21} 系列は最低周波数にピークを持つ。ただし、 AR_{21} の方が高周波領域におけるパワーが大きい。一方で、 AR_{22} 系列は前者2つの系列よりは高い周波数領域にピークを持つ。乱数系列は、全周波数領域にほぼ一樣のパワーを持つ。

3. 計算理論のまとめ

以上、用いた系列の分析から、実験課題の計算理論をまとめると次のことが言える。

視標系列には、ノイズ項が含まれているので（表 4.1），最適な予測をした場合、ヒトの予測誤差はノイズの標準偏差程度となる。また、ヒトが予測しない場合、つまり常に特定の位置だけを見続ける場合には、予測誤差は視標系列の標準偏差の大きさとなる。ここで、ノイズ項の標準偏差は視標系列の半分程度であるから、最適予測した場合と予測しない場合を予測誤差によって切り分けられる実験設計となっている。また、周期的なモデルで AR 系列を予測した場合にも、予測誤差が最適な予測より大きいことがわかり、非周期的な予測モデルで予測する必要のある視標系列であることがわかった。これにより、行動実験では、予測誤差を調べることで、ヒトが非周期的系列をどの程度予測できるか定量的に評価することができる。AR 次数など系列の特性も、500 試行のサンプルから良く推定できた。

第5章 系列予測行動結果

本章では，実験参加者の行動について分析した．主だった分析は系列分析と同様に，実験参加者の予測誤差，AR 次数，AR パラメータである．

予測誤差の解析について，以下の2つの解析を行った．

1. 予測誤差

1.1 対照実験との比較

各タスクでの予測誤差の標準偏差を求め， AR_1 タスク， AR_{21} タスク， AR_{22} タスクについて実験者毎に予測できているかどうか， $SD_{ar}\{E(T)\}$ が $SD_{rnd}\{E(T)\}$ よりも小さくなることを有意水準 5 % で等分散性検定によって調べた．

1.2 学習について

実験参加者が課題遂行中に AR ダイナミクスを学習している様子を観測するために，AR タスクにおける最初 150 試行と最後 150 試行の予測誤差を求めて比較した． AR_1 タスクでは，実験参加者 NN，TY で最後 150 試行の RMSE が有意に小さく ($p < 0.01$, $p < 0.05$)， AR_{21} タスクでは，実験参加者 KY，NN で最後 150 試行の RMSE が有意に小さかった ($p < 0.05$, $p < 0.05$, 図 5.1)． AR_{22} タスクでは，最初 150 試行と最後 150 試行で有意に差のある実験参加者はいなかった．

AR_1 タスクと rnd_1 タスクの予測誤差を解析した結果， $SD_{ar}\{E(T)\}$ が $SD_{rnd}\{E(T)\}$ との比較において，全ての実験参加者で有意に小さかった ($p < 0.01$, 図 5.3)．

同様に， AR_{21} タスクと rnd_{21} タスク， AR_{22} タスクと rnd_{22} タスクを比較した結果も，全ての実験参加者で有意に $SD_{ar}\{E(T)\}$ が小さかった ($p < 0.01$, 図 5.3)．

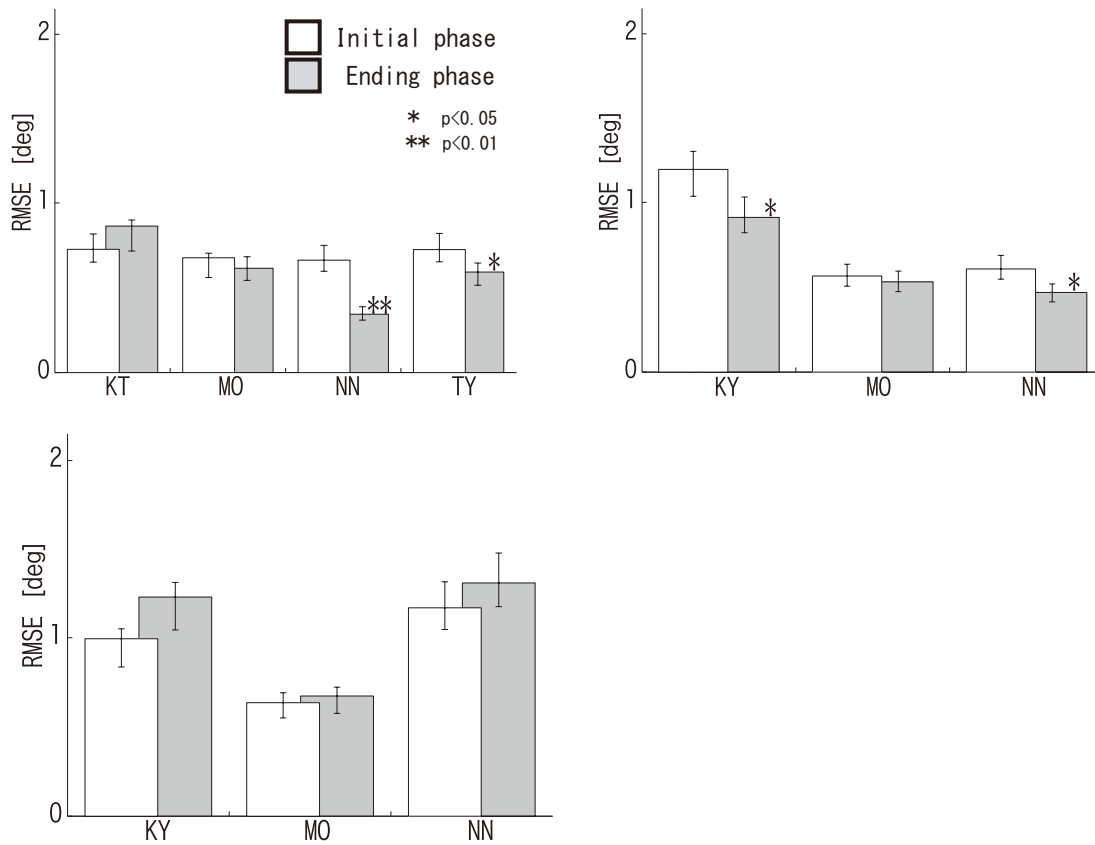


図 5.1: 実験参加者の予測誤差：AR タスクの前 150 試行と後 150 試行の比較.
 AR タスクにおいて、最初 150 試行（白）と最後 150 試行（灰色）における予測誤差を比較した（左上図， AR_1 タスク；右上図， AR_{21} タスク；左下図， AR_{22} タスク）。 AR_1 タスクでは、実験参加者 NN と TY で有意な差があり， AR_{21} タスクでは，KY と NN で有意な差があった。

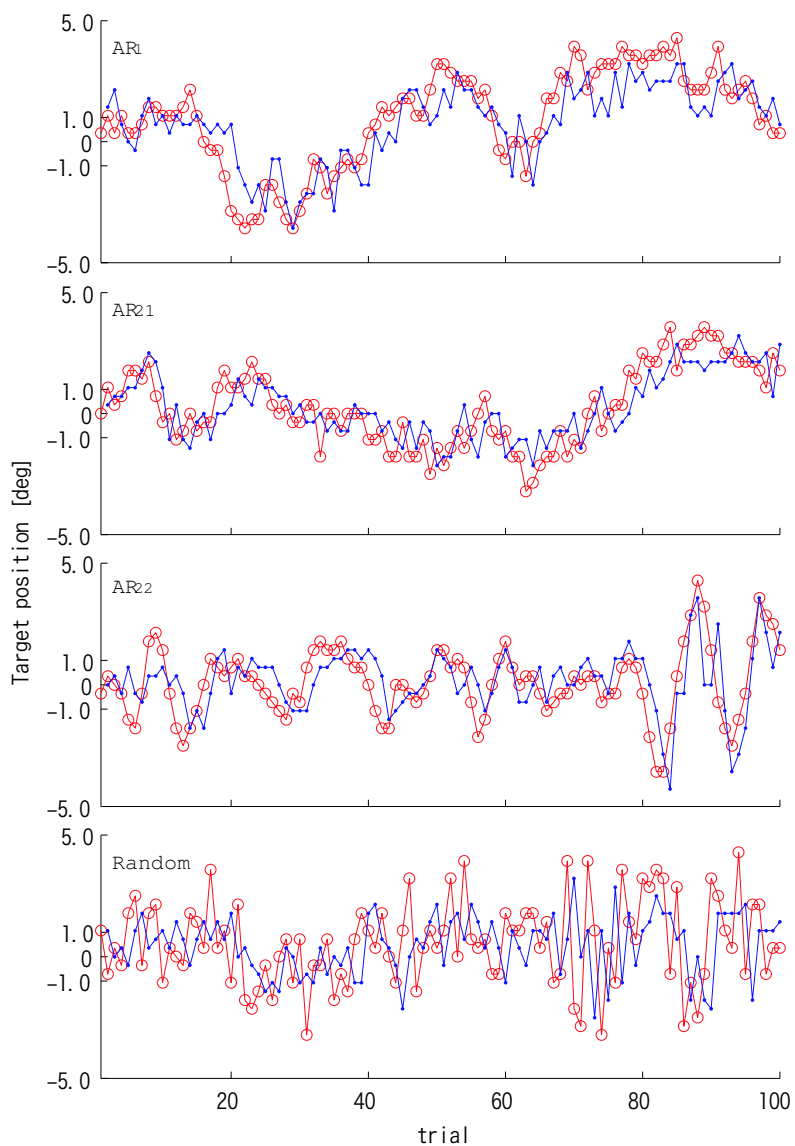


図 5.2: 実験参加者の予測系列の 1 例.

上から AR_1 タスク, AR_{21} タスク, AR_{22} タスク, rnd_1 タスクにおける視標系列 (赤線) と実験参加者 NN の予測系列 (青線) である. 他の実験参加者の結果など本研究で行った全タスクの結果 width=0.8 は, この章の最後に図示した.

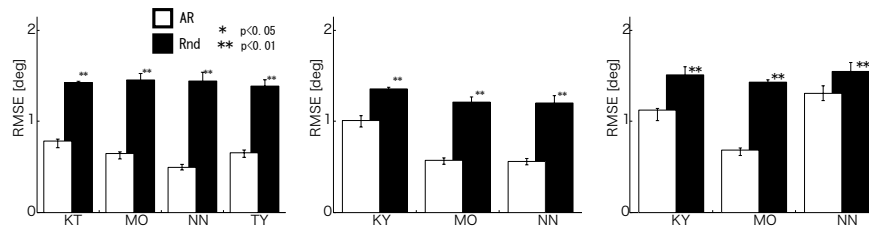


図 5.3: 実験参加者の予測誤差・AR タスクと乱数タスク比較.

左から AR_1 タスクと rnd_1 の比較, AR_{21} 系列と rnd_{21} の比較, AR_{22} 系列と rnd_{22} の比較である. 横軸が実験参加者, 縦軸が RMSE である. 全ての比較について, 全ての実験参加者で AR タスクでの予測誤差が乱数タスクのそれよりも有意に小さい値であった.

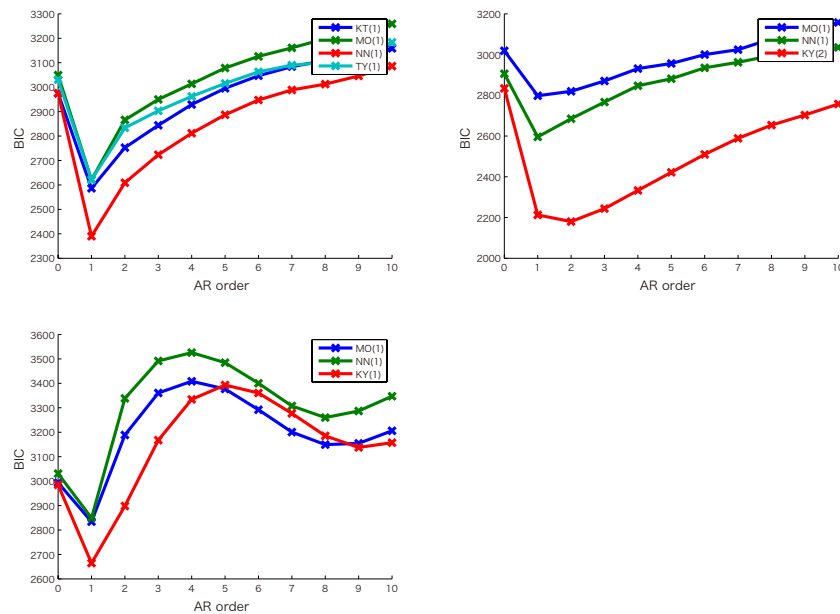


図 5.4: 実験参加者の AR 次数推定: AR タスク.

AR タスクで実験者が行った予測の系列について BIC の値を次数毎に求めた. 左上が AR_1 タスク, 右上が AR_{21} タスク, 左下が AR_{22} タスクについてである. 横軸が AR 次数, 縦軸が BIC の値であり, 各実験参加者を色を変えて表示した. 図の凡例において実験参加者の横の括弧で囲われた数字は BIC 値が最小となった AR 次数である.

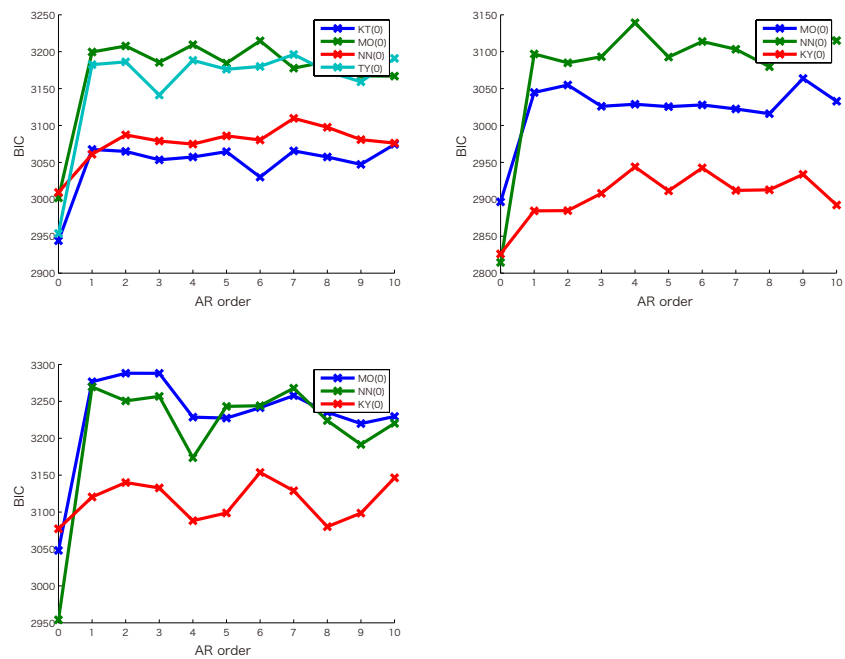


図 5.5: 実験参加者の AR 次数推定・AR タスク.

乱数 タスクで実験者が行った予測の系列について BIC の値を次数毎に求めた. 左上が rnd_1 タスク, 右上が rnd_{21} タスク, 左下が rnd_{22} タスクについてである. 横軸が AR 次数, 縦軸が BIC の値であり, 各実験参加者を色を変えて表示した. 図の凡例において実験参加者の横の括弧で囲われた数字は BIC 値が最小となった AR 次数である. 全ての乱数タスク, 全ての実験参加者について BIC の最小値は 0 次の場合であった.

表 5.1: 実験参加者の予測行動の AR パラメータ. 上から AR_1 タスク, AR_{21} タスク, AR_{22} タスクの結果である. 各段の一番上が系列の AR パラメータである. 各実験参加者については, BIC を用いて推定された次数の AR モデルだとした場合の AR パラメータを推定した (5.4). 2 次 AR モデルだと推定された AR_{21} タスクにおける KY についても, 2 次の項の AR パラメータは非常に小さい値であり, 1 次 AR モデルに近いと言える.

AR1	$a_1 = 0.89$	$a_2 = 0.0$
KT	0.74	-
MO	0.85	-
NN	0.81	-
TY	0.82	-
AR21	$a_1 = 0.58$	$a_2 = 0.31$
KY	0.82	-0.05
MO	0.87	-
NN	0.75	-
AR22	$a_1 = 1.29$	$a_2 = -0.68$
KY	0.70	-
MO	0.86	-
NN	0.66	-

2. 実験参加者の予測行動系列解析

2.1 実験参加者の予測行動の AR 次数

それぞれの AR タスクにおける AR 次数を BIC をもとに同定した結果では, AR_1 タスクでは全ての実験参加者が 1 次と同定された (図 5.4). AR_{21} タスクでは MO と NN が 1 次, KY が 2 次と同定された. AR_{22} タスクではすべての実験参加者が 1 次と同定された. また, 乱数タスクでは, すべての実験参加者が 0 次と同定された (図 5.5).

2.2 実験参加者の予測行動の AR パラメータ

AR パラメータの同定結果は, AR_1 タスクでは, 実験参加者 KT で $a_1 = 0.75$, NN で $a_1 = 0.81$, MO で $a_1 = 0.85$, TY で $a_1 = 0.82$ であった (表 5.1). AR_{21} タスクでは, 実験参加者 KY で $a_1 = -0.05, a_2 = 0.82$, NN で $a_1 = 0.75$, MO で $a_1 = 0.87$ であった. AR_{22} タスクでは, 実験参加者 KY で $a_1 = 0.70$, NN で $a_1 = 0.66$, MO で $a_1 = 0.86$ であった.

2.2.1 特異な予測モデルであることの有無

表 5.2: 特異的な戦略の有無. 全ての実験参加者について 1 次 AR モデルだとして AR パラメータを求め、その 95 % 信頼区間をあわせて示した. 仮に前試行の位置を予測位置とした場合、実験参加者の予測行動は、誤った予測をしているにもかかわらず、1 次 AR モデルと推定される. ただし、この場合は、1 次の AR パラメータが 1.0 となる. 実験参加者の AR パラメータは 1.0 とは有意に差があり、そのような特異的な予測戦略を用いていないことがわかった.

	KT	MO	NN	TY
AR_1	$0.74 \pm 0.2 \times 10^{-2}$	$0.85 \pm 0.2 \times 10^{-2}$	$0.81 \pm 0.2 \times 10^{-2}$	$0.82 \pm 0.2 \times 10^{-2}$
	KY	MO	NN	
AR_{21}	$0.78 \pm 0.4 \times 10^{-2}$	$0.87 \pm 0.2 \times 10^{-2}$	$0.75 \pm 0.2 \times 10^{-2}$	
AR_{22}	$0.70 \pm 0.3 \times 10^{-2}$	$0.86 \pm 0.1 \times 10^{-2}$	$0.66 \pm 0.3 \times 10^{-2}$	

ダイナミクスを用いず、特異的な戦略をとることで、実験参加者の予測の AR 次数が 1 次と同定される場合が想定された. それは、実験参加者が、1 試行前の視標位置を予測位置とする場合である. その場合、予測は最適ではなく、特異的な戦略をとっていたとすれば、1 次 AR パラメータは 1.0 となるが、予測行動は視標系列と 1 次 AR 系列だと同定される. このような特異的な戦略をとっていないか、全ての実験参加者の予測行動をで 1 次 AR モデルだとして、その 1 次 AR パラメータを求めた. その結果、AR パラメータの値は、すべての実験参加者で、1.0 の場合と有意な差があった (図 5.2).

3. 行動予測

AR タスクにおける前半 250 試行の実験参加者の予測系列から AR パラメータを同定し、後半 250 試行の実験参加者が予測行動を予測した. 実験者の行動予測を $P'(T)$ 、実験参加者の予測行動を $P(T)$ とし、その予測誤差を $E'(T) = P(T) - P'(T)$ とし、評価した. ここでの予測誤差 $E'(T)$ を、前述した予測誤差 $E'(T)$ と区別するために、行動予測誤差と呼ぶ.

実験参加者の AR 次数推定の結果、AR タスクではほとんどの実験参加者が 1 次であると推定されたこと、2 次と推定された 1 名も 2 次の項の AR パラメータはごく小さな値であることから、実験参加者の予測モデルは 1 次 AR であるとして、行動予測を行った. 行動予測誤差の平均は、 AR_1 タスクでは KT で 0.69 deg, MO で 0.51 deg, NN で 0.32 deg, TY で 0.50 deg であり、 AR_{21} タスクでは KY で 0.77 deg, MO で 0.34 deg, NN で 0.49 deg, AR_{22} タスクでは KY で 0.82 deg, MO で 0.37 deg, NN で 0.73 deg であった.

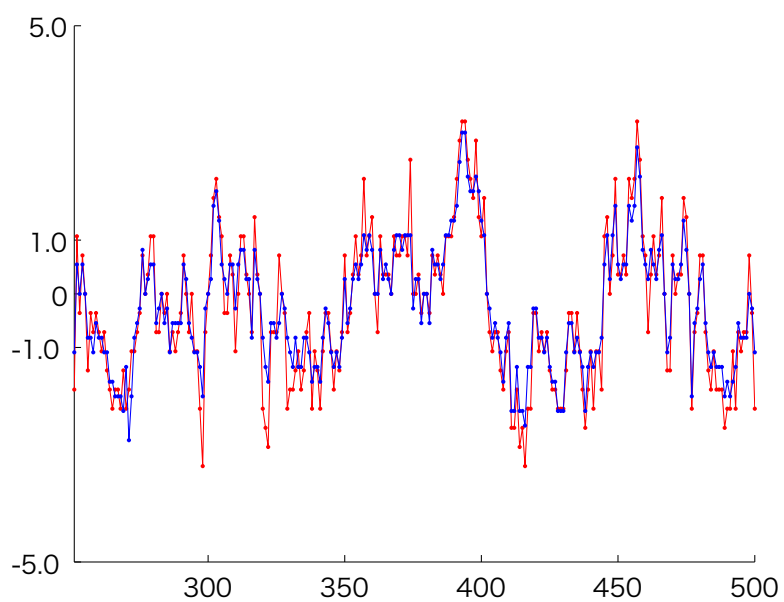


図 5.6: 実験参加者行動の予測の 1 例, AR_1 タスク, NN. 赤線が実験参加者の行動であり, 青線が行動予測である. 他の実験参加者については, 補足に載せた.

第6章 系列予測についての議論

本研究ではヒトがダイナミクスを持つ視標時系列に対し、そのダイナミクスを用いて予測をすることができかどうか調べた。

1. ダイナミクスをもつ系列の予測

まず、AR タスクと乱数タスクでの予測誤差の標準偏差 $SD\{E(T)\}$ を比較することにより、実験参加者の予測性能を調べた。1 次 AR タスク、および対となる乱数タスクの両時系列は静的統計量が等しいが予測可能性の有無によって予測性能に差があった。AR タスクにおける予測誤差の標準偏差 $SD_{ar}\{E(T)\}$ が乱数タスクにおける予測誤差の標準偏差 $SD_{rnd}\{E(T)\}$ よりも有意に小さかった。また、AR タスクにおける予測誤差の標準偏差 $SD_{ar}\{E(T)\}$ が理論的な予測可能性の基準となる視標時系列の標準偏差 $SD_{ar}\{S(T)\}$ よりも小さく、最適予測の基準となるシステムノイズの標準偏差 σ_E に近いことから、1 次 AR タスクを行った実験参加者全員が 1 次 AR 過程に対して準最適な予測性能を持つ事が示唆された。同様に、2 次 AR タスクにおいても行った実験参加者全員が 2 次 AR 過程を準最適な予測をしていることが示唆された。

ただし、AR タスクにおける予測誤差の標準偏差 $SD_{ar}\{E(T)\}$ が最適予測の基準となる σ_E とは有意な差があることや、モデル同定分析結果で実験参加者の予測時系列が視標時系列の生成モデルとは異なっていることから、最適予測を行っていたとは言えない。ヒトが AR モデルで予測を行っていたかはわからない。ヒトの予測モデルは、AR モデルを獲得していたが AR 係数の異なるモデルである可能性、AR モデルではない予測モデルである可能性、もしくはいくつかのモデルの複合による予測モデルである可能性が挙げられる。

2. ヒトの予測行動の特性

予測モデルの解析から準最適な予測性能は、単に視標位置記憶を用いた予測による結果ではないことがわかった。AR タスクでは 1 人を除いて全て実験参加者の予測行動で 1 次 AR だと同定された。また、2 次 AR と同定された 1 人について

ても、2次 AR パラメータは小さい値であり、実質的には1次 AR モデルであると考えてもよい。本実験課題で実験参加者は、過去1試行の視標位置にもとづいて予測したと考えられる。

先行研究に、サッカーのリアクションタイムが過去1－3試行に影響されるという研究や、円滑性眼球運動の速度が過去2試行に影響されるという研究が報告されている。ヒトがどれだけ過去に遡った情報に基づいて予測するかは、記憶容量とも密接に関わる。本実験課題では、2次 AR タスクは2試行前までに基づいて予測することが最適であるにもかかわらず、実験参加者の予測行動の AR 次数が1次と同定されたが、これは視覚刺激の提示として過去の試行を記憶しにくかった可能性がある。1試行前の視標位置は直前であるが、2試行前の視標位置は4秒以上前の情報である。今後、離散的な系列において、試行間の時間がどのように影響するかを調べる必要がある。

乱数タスクの予測時系列を調べてみると、予測誤差最小という点では常に0 mm の位置を予測することが乱数タスクにおける最適予測の1つと定義でき、逐次最小二乗法を用いた計算機実験でもそのような予測時系列が観測される。一方で実験参加者の予測系列は位置0 mm を予測し続けるのではなく、ランダムな挙動を示した。AR タスクと乱数タスクで予測モデルが異なり、視標系列のダイナミクスの有無を知覚して予測モデルを変化させている事が示唆された。

3. 行動予測の妥当性

行動予測誤差は、実験参加者の行動を周期的なモデルで正弦波モデルだとして予測した場合よりも小さく、比較してARモデルで良く予測できたと言える。

行動予測誤差の結果は予測誤差の結果と相関しており、予測誤差の小さな実験参加者では行動予測誤差も小さかった。行動予測誤差が小さいということは、前半と後半での予測行動のダイナミクスが変化していないことを示しており、実験参加者が安定的に1つのダイナミクスで予測していることを示す指標でもある。視標系列を良く予測するためには、視標系列のダイナミクスを把握すると同時に、その把握したダイナミクスを確信して予測する必要がある。逆に言えば、確信できず探索的に予測行動のダイナミクスを変化させている場合には、予測誤差が大きくなる。予測誤差が小さく良い予測が出来ている実験参加者ほど、安定的に予測していることがわかった。逆に、予測誤差が大きかった実験参加者は、視標系列のダイナミクスを十分に確信して同定できておらず、それが行動予測誤差の大きさに反映されているとも考えられる。

ただし、今回の実験課題では、予測誤差、行動予測誤差に反映されている事として、予測能力だけでなく、学習の要因が考えられる。今回の実験課題では、500

試行内では、実験参加者が学習していたか、さらに学習がすすむのかについて確定的なことは得られなかった。また異なる AR 次数においても同じ 1 次 AR 次数であると同定されたため、実験参加者が視標系列のダイナミクスに適応的に行動予測のダイナミクスを変えていることもわからなかった。今後、視標系列のダイナミクスが、今回よりも明確に異なる視標系列を用いて、実験参加者の予測行動のダイナミクスが適応的に変化していることを示せる実験設計を行う必要がある。

また、行動予測誤差は、周期的なモデルをもちいて予測した場合よりは良かったが、評価方法についてはさらに検討の余地がある。今回は視標系列の生成モデルである AR モデルでヒトの行動も予測したが、他の非周期的モデルで行動を予測した場合の行動予測誤差を算出して比較することも、ヒトの予測特性を調べるために重要だろうと考えている。

fMRI計測を用いた顔知覚の研究

第7章 顔知覚の研究背景と目的

1. 顔知覚

私たちは社会生活の中で多くの人に出会い、様々な情報を顔から推測している。顔を見ることは非常に身近な行為であり、学術的にも様々な分野が顔を対象に研究を行っている。一概に顔知覚の研究といっても、様々な観点からのアプローチがある。例えば、そもそもどのような図形が顔と知覚されるかという研究や感情を推測するのに必要な表情の研究が行われている。

1.1 顔知覚の計算理論；コンピュータビジョンの先行研究

顔に限らず、視覚的物体認識の計算理論上の問題として、不変性が挙げられる。不変性は、顔や物体を認識するさいに照明や見えの角度などが変化するなど、網膜からの入力像が変化しても、同一物体として認識する必要があるということである。これは入力像と記憶している像の単なるテンプレートマッチングを行うような情報表現では、視覚認識を実現できないことを示している。

コンピュータビジョンの顔画像認識の分野では、1987年に Sirovich と Kirby は、複数人の顔画像の固有ベクトルで顔画像の特徴が表現できることを示した [59]。これは、顔画像が、特に正面顔について考えると、どの人の顔も目や鼻や口など構造的には同じ配置にあり、画像間には強い相関があることや、額や頬など視覚的に均一な面積も多いからであると考察できる。その後、1991年に Turk と Pentland は、多数人の顔画像について主成分分析 (Principal Component Analysis; PCA) をおこない、求めた固有ベクトル空間をもと判別モデルを作成し、人物のクラスに分類する実験を行った [65]。実験では、16人の顔についてのさまざまな条件で撮像した 2500 の顔画像からどの人物であるかを推定する実験を行い、照明を変化させただけの条件では 96 % の正解率であった。また、彼らは固有ベクトルに相当する顔画像を固有顔 (Eigen-Face) と呼んだ。このモデルは、平均顔と固有顔という複数の顔テンプレートを持ち、その線形和で表現されるモデルであり、もとの画像空間から考えると情報圧縮されている。情報処理資源がきわめて限られた脳における情報表現として、適したモデルだと考えられる。

1.2 脳の顔知覚の情報表現：生理学，心理学の先行研究

心理物理学の分野で 1986 年に Bruce と Young により提案された顔認識の情報処理過程について，脳で対応する領野を示したモデルが報告されている [7, 28].

サルを対象にした電気生理学研究では，顔の部分である目や鼻や口などに選択的に活動する細胞であったり，その大きさに応じて活動が変化する細胞が見つまっている [69, 17]. また，ヒトの顔であるかサルの顔であるかを識別する大局的な情報処理が行われてから，さらにどのサルの顔か，どのヒトの顔か，どの表情であるかなどの局所的情報が処理される [62].

fMRI 研究では，紡錘状回（Fusiform）に顔に選択的な活動があると報告され，Fusiform Face Area（FFA）と呼ばれている [32, 24, 60]. 物体なども色や形が統合された視覚刺激であるが，顔は物体が表現されている領野とは異なる．FFA は単に入力された顔刺激を情報処理するだけでなく，顔知覚に関わっているとされており，顔画像の視覚刺激入力ない場合でも顔を知覚すれば FFA が活動する [70].

さらに情報表現に踏み込んだ研究では，形や肌質が脳活動に反映されていることが報告されている．[29] また，顔知覚空間を仮定して連続的に任意の顔画像が作成できるモデルを用いた研究では，見せた顔画像間の距離によって脳活動が変化したという報告がされている．[36, 30].

2. 本研究の目的

本研究では，視覚的に個人を同定する時の顔知覚について調査した．脳の高次視覚領野である下側頭葉における顔知覚のモデルを考え，そのモデルによって顔を知覚している際の下側頭葉の脳活動が説明できる事を示す．

心理学では，1996 年に顔の類似性が PCA 空間と相関するという報告がされている [26, 49]. これらの研究では 2 次元程度の低次元の PCA 空間のみで議論しており，また内観に基づく調査である．明示的ではないが，PCA 基底に基づくかお画像を作成して行動課題を行った研究もあるが，本研究とは，ことなり適応知覚実験であった．本研究で行った識別課題とあわせて考察する必要がある．本研究では実験課題遂行中の脳活動を fMRI により計測した．実験課題では，実験参加者は顔画像を見て顔を知覚し，次に提示された 2 つの異なる顔のどちらが知覚した顔に近いかを識別した．信号検出理論を基にした顔識別モデルを用いて，被験者の顔識別行動を良く予測できることで，モデルの妥当性を示す．

また，顔知覚モデルによって下側頭葉の脳活動を説明するために，脳活動から知覚している顔の顔知覚座標を推定することを試みた．

第8章 実験方法

1. 顔識別課題

実験課題では顔画像を 1000 ms 間表示し、500 ms のブランクの後に、選択肢となる 2 つの顔画像を同時に表示した。被験者は最初に表示された顔画像を記憶しておき（顔知覚期間）、選択肢となる顔画像が表示されたらば、ただちに記憶していた顔が選択肢顔画像のどちらに近いかをボタン押しによって回答する。これを 1 試行とした。

2. 顔知覚空間

本研究では、顔を見たときに選択的に活動を示す高次視覚領野（顔領野）における顔知覚表現を主成分分析（Principal Component Analysis, PCA）の基底で構成される顔知覚空間と考え、また、この顔知覚空間モデルを生成モデルとして fMRI 実験で用いる顔画像刺激を作成した。

20 人の原顔画像 $F_i^{original}$ ($i = 1 : 20$) の 61 の対応点座標 S と 20 人の空間的平均顔の輝度 T について、それぞれを PCA をかけて 19 次元の基底ベクトル $e_F = (e^S, e^T) = (e_1^S, \dots, e_M^S, e_1^T, \dots, e_M^T)$ を求めた。それぞれの基底空間は 19 次元で 99 % 以上の累積寄与率となった。輝度値から求めた基底ベクトル 19 次元を顔知覚空間における最初の 19 次元とし、対応点から求めた基底ベクトル 19 次元を顔知覚空間における 20-38 次元目として、38 次元の顔知覚空間を作成した。

この様に表現された顔画像（顔知覚）表現では、個々の顔知覚はその空間の一点として与えられる。顔知覚空間は、連続空間であり、任意の顔を表現することが可能である。入力された視覚情報（網膜像、低次視覚野での表現）から情報縮約された情報表現により顔知覚が実現される。2 つの顔を識別するために（もしくは複数の顔をクラス分類するために）最適な情報表現ではなく、顔を同定するために最適化された情報表現である。また、PCA 基底空間は線形空間であり、ニューラルネットワークで実現可能であることが理論的にわかっている。

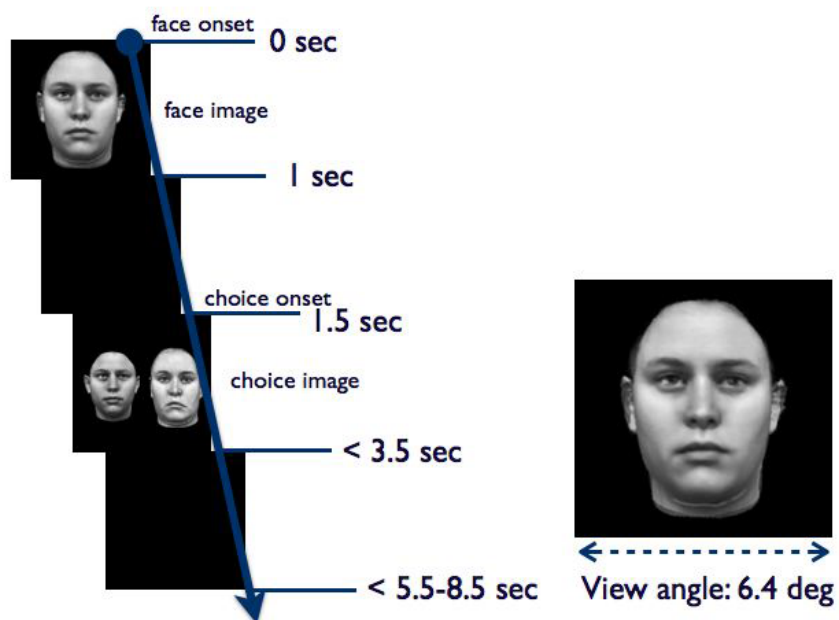


図 8.1: 顔識別課題.

顔画像を 1000 ms 間表示し、500 ms のブランクの後に、選択肢となる 2 つの顔画像を同時に表示した。被験者は最初に表示された顔画像を見て記憶しておき、選択肢となる顔画像が表示されたらば、ただちに記憶していた顔が選択肢顔画像のどちらに近いかをボタン押しによって回答する。これを 1 試行とし、63 人分の顔画像を 3 回ずつ、ランダムな順序で、計 189 試行を行った。最初の顔画像は視野角 6.4 deg 以内となるようにした (右図)

3. 顔識別モデル

知覚した顔を異なる2つ顔のどちらかに識別する行動について説明する。

顔画像 F_j を見たときに顔 F_u を知覚する確率を $p_p(u|W_j)$ とする。この確率は平均 $\mu_u = W_u$ 、分散 $\sigma_u = \sigma$ の正規分布で与えられ、この確率を知覚確率と呼ぶ。同様に、 F_j を見たときに F_v と知覚する確率 $p_p(v|W_j)$ は平均 $\mu_v = W_v$ 、分散 $\sigma_v = \sigma (= \sigma_u)$ の正規分布で与えられる。知覚確率の分散は被験者により異なり、分散が大きい被験者ほど識別確率の傾きが全体的に小さくなる。

F_u と F_v を選択肢とし、任意の顔 F_j を見て似ている顔を回答する強制2択課題を課した場合、被験者の識別行動は、 F_v を選択する確率 $P_d(v|W_j)$ として表現し、 $P_d(v|W_j) = p(v|W_j)/(p(u|W_j) + p(v|W_j))$ と与えられる。

顔識別モデルに関する補足

知覚確率の分散は被験者により異なり、分散が大きい被験者ほど識別確率の傾きが全体的に小さくなる。今回は人種効果（見慣れない人種の顔の同定や識別が困難となる現象）を考えないので、知覚確率の分散は全ての顔で等分散であると仮定する。人種効果を考えた場合は、顔によって分散が異なる場合を考える必要があるかもしれない。ただし、その場合でも分散の変化は連続的であり、近い距離にある顔同士では大きく変わらないと推測する。

一般的に知覚が確率分布として表現される起源は、顔を表現している神経活動が持つゆらぎ、視覚情報処理過程で生じる影響、高次領野からのトップダウン信号の影響などが考えられるが、ここではその起源を特には仮定しない。

4. 実験設定の詳細

4.1 原顔画像

顔画像のデータベースは Volker Blanz (Max-Planck Center for Visual Computing and Communication) から取得した。これはカラーの3次元画像のデータベースであるが、本研究では正面を向いている2次元顔画像を20人分をグレースケールにして用いた（図8.3）。この20人の顔画像を原顔画像 $F_i^{original} (i = 1 : 20)$ とする。画像は正方形で一辺の大きさは256 pixelであり、全65536 pixelとなる。各pixelは0 - 255の輝度値を取る。また、異なる顔同士において顔の輪郭や目尻など顔を構成する要素の空間的一致をとるために、各原顔画像から顔要素配置

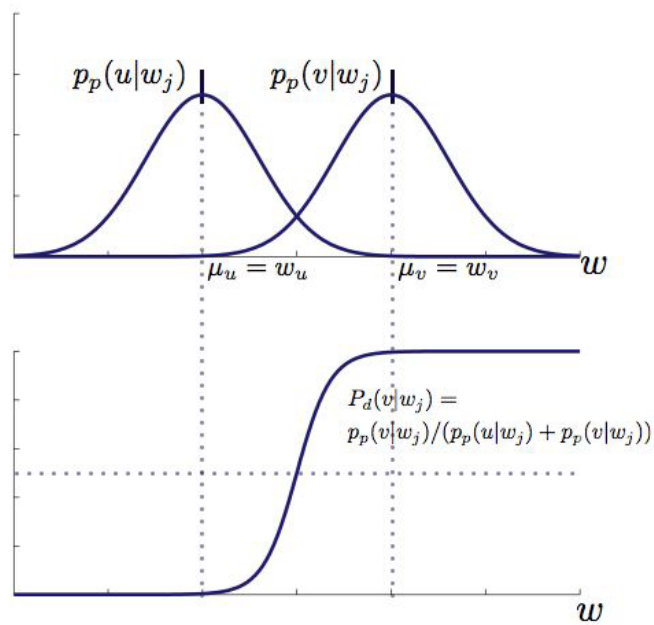


図 8.2: 顔識別モデル.

上図：横軸，顔知覚空間上の座標；縦軸，顔を知覚する確率であり，顔画像 u の顔知覚は顔知覚空間 w 上の座標 w_u を平均とした確率分布 $P_p(u|w_j)$ (u の顔知覚確率) で表現される．つまり，ヒトの顔知覚には揺らぎがあると仮定しており，顔画像 j を見 [htb] た場合にも，顔 u を知覚する場合があります，その確率は $P_p(u|w_j)$ と表される．下図：顔 u と v を選択肢とした強制二択課題において，ある顔画像 j を見た場合に v を選択する識別確率 $P_d(v|w_j)$ は， u と v の顔知覚確率から計算されシグモイド関数となる．

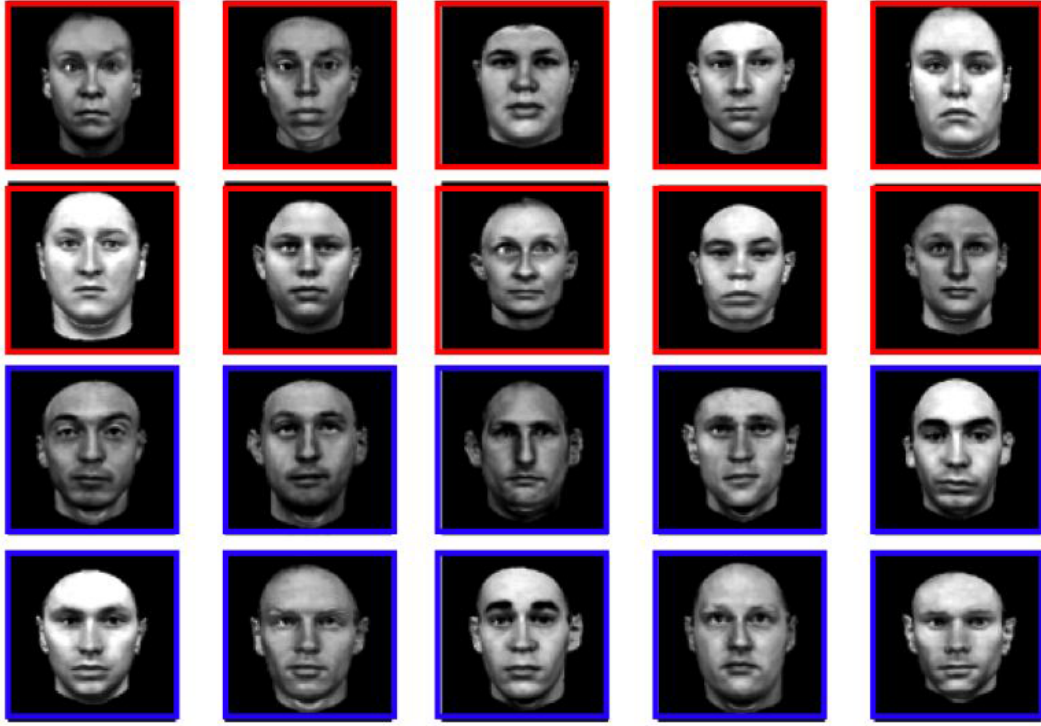


図 8.3: 20 の原顔画像.

顔知覚空間および平均顔画像の作成にもちいた 20 の原顔画像である. 赤枠で囲われた上 2 段 10 の顔画像が女性の顔画像であり, 下 2 段 10 の顔画像は男性の顔画像である. 各顔画像は 1 辺の大きさ 256 pixel 正方形であり, 各 pixel は 0-255 の輝度値をとるグレースケール画像である.

の目印となる 121 点の対応点座標を取得し, そのうちの 61 点を実際に用いた (図 8.4) .

つまり, 20 の原顔画像のデータは

$$F_i^{original} = (S_i, T_i),$$

$$\text{輝度 (Texture)} : T_i = (t_{i,1}, \dots, t_{i,65536}),$$

$$\text{形状 (Shape)} : S_i = (s_{i,1}, \dots, s_{i,61}), s_{i,j} = (x_{i,j}, y_{i,j}),$$

のベクトルとして表現される.

4.2 平均顔画像の作成

61 の対応点 S_i について 20 人原顔画像の平均を求め, これを平均顔画像の対応点座標とした. さらに各原顔画像について, 対応点を三角形の頂点とするドロネー三角形分割を行い, 分割された各ドロネー三角形を局所領域とて, 原顔画像



図 8.4: 顔画像の対応点. 顔知覚空間を作成する際に顔の形状の情報を付加した. その際に各原顔画像から顔要素配置の目印となる 121 点の対応点座標 (緑点) を取得し, そのうち 61 点を用いた.

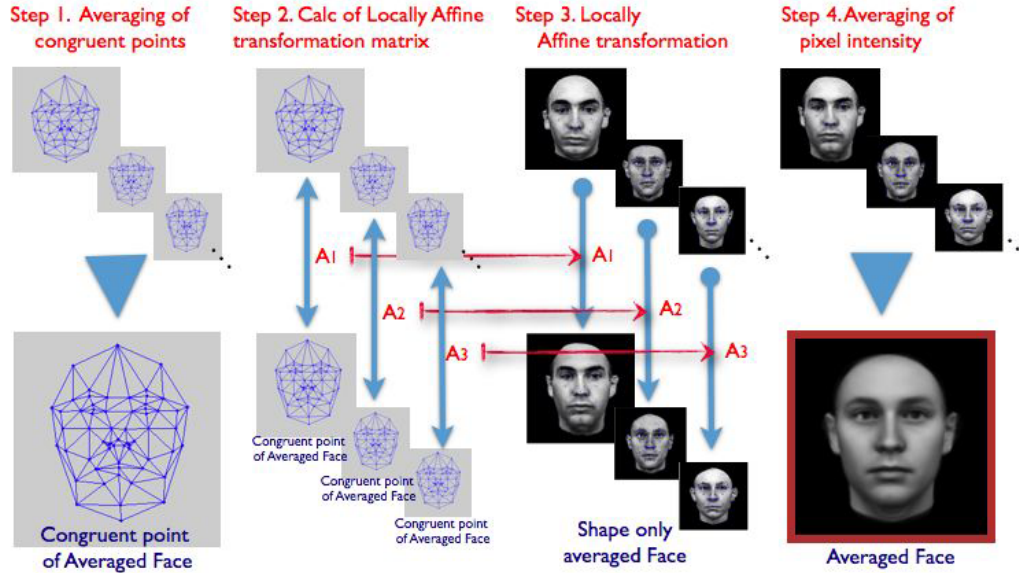


図 8.5: 顔知覚空間における原点となる平均顔作成アルゴリズム.
 左から順に平均顔作成のアルゴリズムを示した. まず, 61 の対応点座標について 20 原顔画像の平均を取った (Step 1). 次に各原顔画像の対応点座標から Step 1 で求めた平均対応点座標へのアフィン変換行列を求めた (Step 2). Step 2 で求めたアフィン変換行列によって, それぞれの顔画像を変換し, 20 の構造平均顔画像を作成した (Step 3). 最後に各ピクセルについて, 20 の構造的な平均顔画像の平均を取り, 平均顔画像を作成した (Step 4). 本研究の顔知覚モデルでは, 脳の顔領域では平均顔をもっていると仮定している.

から平均顔画像へアフィン変換 (局所アフィン変換) をおこない, 顔全体と部分の位置や大きさの等しい 20 人分の空間的平均顔を作成した. そして, 作成した 20 人の空間的平均顔の各ピクセル輝度について平均を取り, 平均顔 F^{mean} を作成した.

4.3 変形顔の作成

38 の基底ベクトルと平均顔画像を用いることにより, 顔知覚空間における任意の座標にある顔画像を作成することが可能である. 顔 F_j の座標 $W_j = (W_j^S, W_j^T) = (w_{j,1}^S, \dots, w_{j,M}^S, w_{j,1}^T, \dots, w_{j,M}^T)$ が得られているとして, 顔画像は次の手順で作成した (図 8.6). まず, 対応点の基底空間座標 $W_j^S = (w_1^S, \dots, w_M^S)$ と対応点基底ベクトル e^S から対応点座標 S_j を求めた. 次に, 平均顔の対応点座標から推定した対応点座標 S_j への局所アフィン変換行列を求めた. ここで, 輝度の主成分得点 $w^T = (w_1^T, \dots, w_M^T)$ と輝度基底ベクトル e^T から各 pixel の輝度を推定し, 対応点

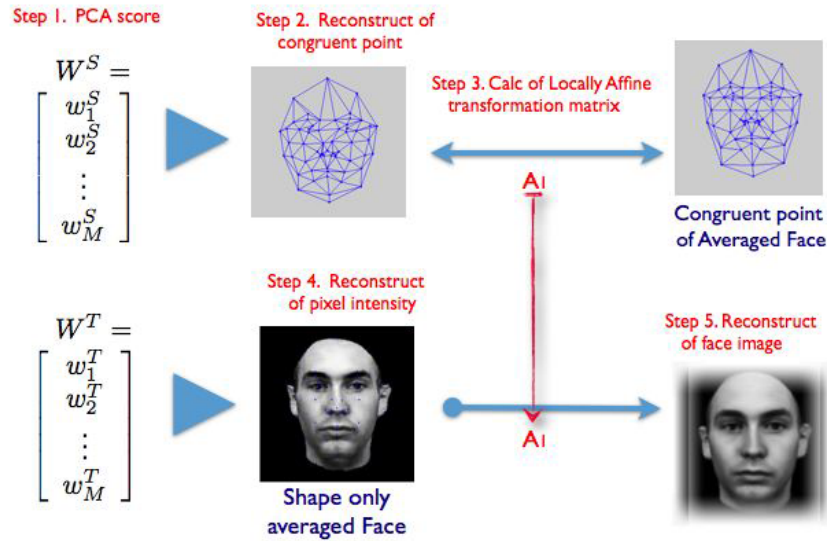


図 8.6: 変形顔作成および顔知覚再構成アルゴリズム.

顔知覚空間上の座標が PCA 得点として与えられており (Step 1), 顔画像を作成 (再構成) することを考える. 座標における形状次元 (Step 1 の上部のベクトル: 全 38 次元の内の 1 – 19 次元) から, 作成する顔画像の対応点座標を求めた (Step 2). 次に, 平均顔の対応点座標から作成する顔画像の対応点座標へのアフィン変換行列を求めた (Step 3). 座標における輝度次元 (Step 1 の下部のベクトル: 全 38 次元の内の 20 – 38 次元) から, 構造的な平均顔画像を求めた (Step 4). Step 3 で求めた変換行列で構造平均顔画像を変換し, 顔知覚空間上での座標が表現する顔画像を作成した (Step 5).

座標が平均と等しい構造平均顔画像を求めた. 最後に, 先に求めた局所アフィン変換行列により, 構造平均顔画像の対応点座標を F_j の対応点座標へと変換し, 顔画像 F_j を作成した.

4.4 実験で用いた顔画像

実験で用いる顔画像刺激は次のように作成した. 原顔画像から適当な 2 つの顔画像を選び, PCA 基底で表現される顔知覚空間上で 8 等分し, 原画像も含め 9 人分の顔知覚空間座標を求め, 求めた顔知覚空間座標から, 顔画像を作成した. PCA 基底表現が脳内における顔知覚の情報表現だとすれば, 心理的に等距離で変化する 7 つの中間顔が新たに作成されたことになる. 7 組の原顔画像を選び, 原顔画像も含めて, 全部で 63 の顔画像を作成した (図 8.7). 7 組の 2 つの原画像間の距離は, 顔知覚空間上で等しくなるように選択した (図 12.27). また, 作成した顔は顔知覚空間上において一部に偏らないように設計した (図 8.9)



図 8.7: 実験で用いた顔画像刺激。

実験で用いた 63 の顔画像刺激である。各段の左端および右端の列の顔が、選択肢顔画像に用いた画像であり、2 列目から 8 列目の顔画像は同じ段の左端と右端の中間顔である（以後、1 つの段における 9 つの顔をセットと呼ぶ。）左端列から右端列の顔の距離の 8 等分点に各列の顔があるように設計した。

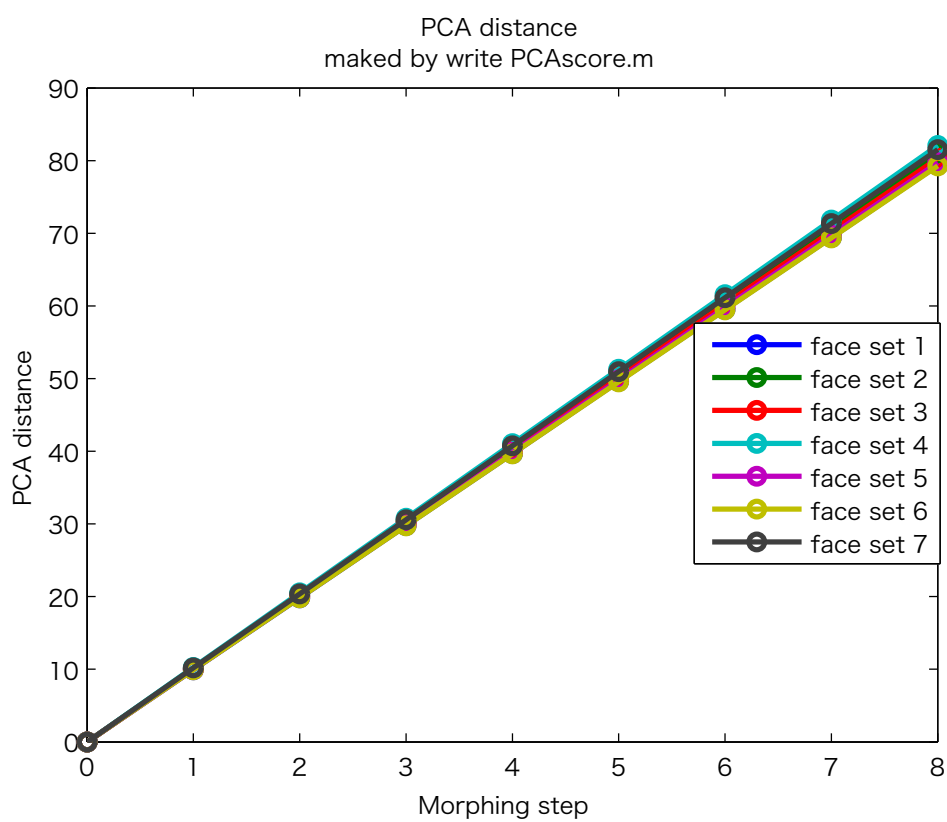


図 8.8: 顔画像の顔知覚空間上での距離。

8.7において、左端の顔から同段の他の顔までの距離を示した。各セットの左端と右端の距離をほぼ等しくなるように設計した。

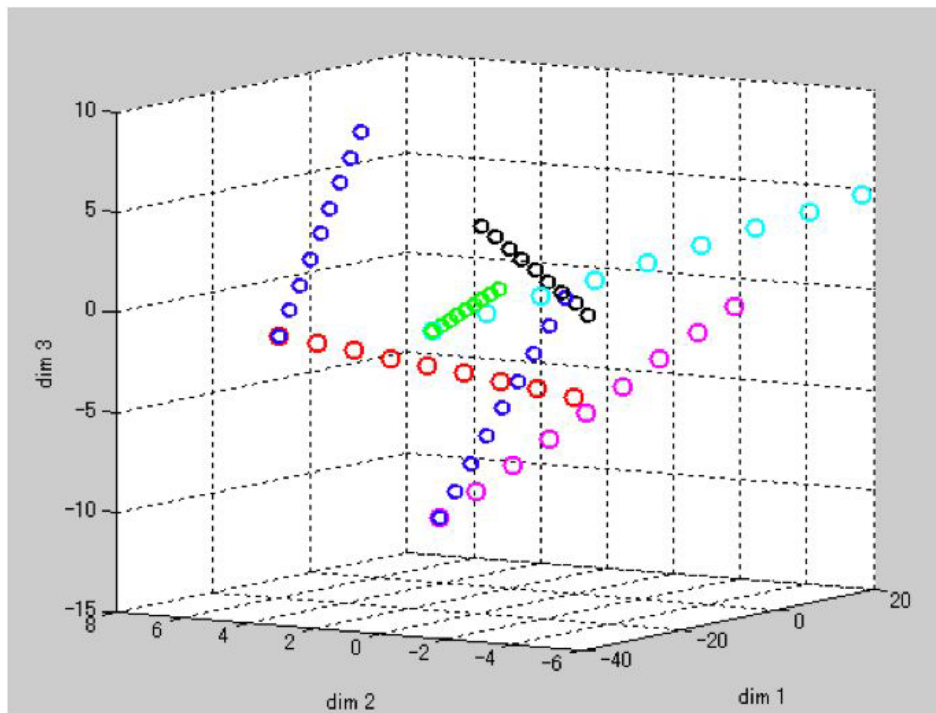


図 8.9: 顔知覚空間上における各顔の位置.

顔知覚空間における最初の3次元空間で各顔の位置を図示した. 同じセットの顔は同じ色で表現した. セット内における各顔は等間隔で並んでいることがわかる. また空間上のさまざまな位置の顔画像を用いたことがわかる.

第9章 顔識別行動結果

以上の実験設計から，仮説としたモデルが正しければ，Morphing Step 0 の顔を見た場合に顔識別確率が0，Morphing Step 8 の顔を見た場合に顔識別確率が1となるシグモイド関数となる．また，Morphing Step 4 の顔を見た場合に顔識別確率が0.5となる．ヒトの行動を解析してみると，顔識別行動モデルと良く一致することがわかった（図 12.18）．

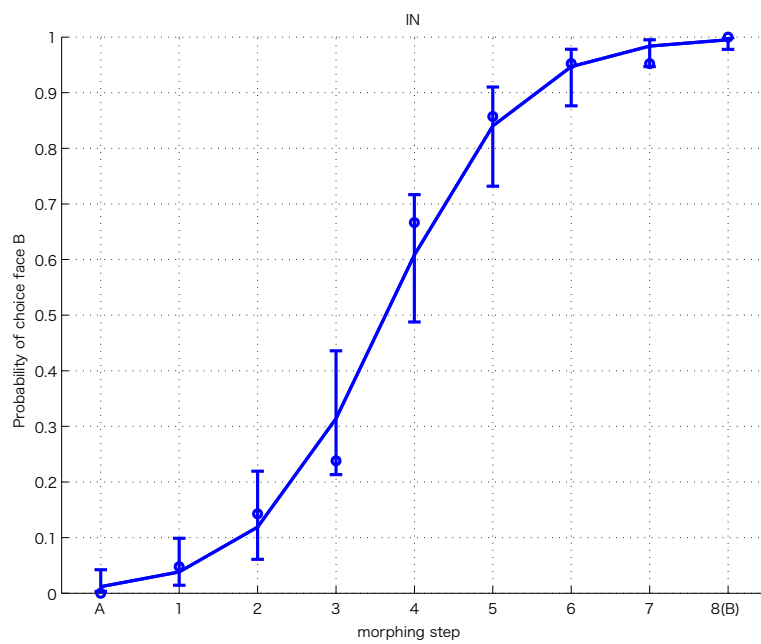


図 9.1: 顔識別行動の一例, 実験参加者 IN. 横軸を Morphing Step とし, 縦軸を Morphing Step 8 の顔を選択したときの顔識別確率とした. 全ての実験参加者の顔識別行動の図は補足を参照.

第10章 脳活動解析

本研究は、行動実験による検証だけではなく、脳活動を解析することによる検証も目指した。特に脳活動の空間的パターンを解析することが近年活発に行われている。この章では、現時点で行った脳活動の解析について議論する。

1. fMRI研究

1.1 fMRI計測

本研究では、fMRI計測を用いた研究を行った。fMRI計測について、後述する議論に必要な知識について説明する。

脳の情報処理は神経活動により行われており、神経活動に伴うエネルギーの消費と脳血流変化と関連していることが知られている [67]。fMRIでは、血液中の酸素濃度に関連した Blood Oxygenation Level Dependent (BOLD) 信号を検出する [43, 44]。具体的には、BOLD 信号は静脈中の還元ヘモグロビン濃度を反映しており、還元ヘモグロビンが血液中に分布している場合には BOLD 信号は低下する。一方で、酸化ヘモグロビンが分布している場合には、BOLD 信号の低下は起こらない。神経活動により血液中の酸化ヘモグロビンが還元されるが、それとともに、神経活動が起きている局所的な血流量が増え、静脈中の酸素が 30 – 50 % 程度増加する。一方で、神経細胞での酸素消費は 5 % 程度の増加にとどまる [16]。その結果、静脈では酸素過剰となり、相対的に還元ヘモグロビン濃度が低下して BOLD 信号が増加する。

実験課題遂行中の BOLD 信号の変化は、通常 1-5 % 程度、大きい場合でも 10 % である。BOLD 信号は、神経活動を起点に増加し、5-7 sec で最大となり、その後 10 sec 程度かけて定常な状態へ低下する。また、fMRI 計測では、一般的に、全脳を 1 回計測する時間は 1-3 sec 程度に設定する。空間的解像度はボクセル (Voxel) と呼ばれる立方体で表され、一辺 1-4mm 程度に設定する。

1 つのボクセルには 10 万個程度の神経細胞が存在すると推測され、1 つのボクセルの BOLD 信号値はそれら神経細胞群の活動の総和と考えられる。神経細胞群の活動を計測した Local Field Potential (LFP) と BOLD 信号が相関する知見がある [38]。また、電気生理学で計測される発火率は計測された神経細胞からの

出力であるが、BOLD 信号は計測された領域ボクセルへの入力信号を反映していると言われている [37]. その他、神経活動と BOLD 信号の関係を数理的に定式化した研究がなされている [18].

以後、簡単のため、脳活動と表記した場合は、fMRI 計測によって得られる BOLD 信号とする.

1.2 これまでの fMRI 研究

これまでの fMRI 研究の多くは、作業仮説として定義した機能が実行されているときに、実行されていないときよりも、偶然ではない程度に高い活動が見られる脳領域 (ボクセル) を、仮説検定により求めていた. しかし、この方法では脳活動の大きさの情報を比べているだけなので、複雑な情報処理については調べることができない. また経験的に、ある情報表現について、さらに細分化して詳細な情報表現を調べようとする、非常に困難が伴う. 例えば、他の物体知覚と比較して顔に選択的に活動する紡錘状回顔領域 (Fusiform Face Area; FFA) が知られているが、ここで顔という情報表現のカテゴリを細分化し、顔 A や顔 B について選択的に活動するボクセルを調べようとしても、個別の顔に選択的な活動がさまざまな領域に現れるか、逆に全く現れないなど、安定した解析結果が得られない. また、個別の顔を表現するヒトに共通の脳領域を調べるために、集団解析をしても有意な活動を示すボクセルは現れないことがほとんどである.

1.3 fMRI デコーディング研究

2005 年に脳活動を解読してヒトが見ている視覚刺激を推定した研究が報告された. Kamitani2005, 2006.

この研究では、実験参加者に 2 種類の縞状視覚刺激を見せ、その時の脳活動から実験参加者が見ている方の刺激を推定した. 脳活動は低次視覚領野を対象としており、数百ボクセルの脳活動データである. 脳活動から視覚刺激を推定するデコーダーはサポートベクターマシーン (Support Vector Machine: SVM) で作成されており、クラス分類を行い、その結果、正解率はチャンスレベルである 50 % を超えた. この研究により、fMRI 計測によって得られる BOLD 信号の脳活動にクラス分類できる程度の情報を持つことが分かった.

その後、SVM を用いたクラス分類 (classification) により脳活動から刺激や行動を推定した研究が多く報告され、これまで fMRI の解析で主流であった差分解析で不可能であった脳情報を判別できることがわかった. 脳活動から刺激や行動を推定するこれらの研究は、デコーディング (Decoding) 研究もしくはブレイン・リーディング (Brain-reading) 研究と呼ぶ.

ただ、クラス分類の方法では、デコーダーが推定する刺激は刺激セットからの選択であり、論理的に考えて刺激セットに含まれない新規刺激については推定できない。これに対して、2008年に、同定（Identification）とよばれる方法で新規120の刺激についても低次視覚野の脳活動から視覚刺激を推定できたデコーディング研究が報告された（Kay2008）。この方法では、まず、視覚刺激から脳活動パターンを予測する符号化モデルを用意し、パラメータの学習をする。学習した符号化モデルに、新規120の視覚刺激を入力し、120の予測脳活動を求めることで、視覚刺激と予測脳活動の対応表を作成する。最後に、実際の脳活動と最も相関する予測脳活動を同定し、先の対応表から同定した予測脳活動の対となる視覚刺激を求め、推定視覚刺激とした。その結果、新規120の視覚刺激について、チャンスレベル（ $1/120 = 0.83\%$ ）を超える正解率を可能にした（正解率は、1回の脳活動を使った場合に32%、51%であり、複数回の脳活動を用いた場合には72%、92%であった）。

また、Kayらの方法では、脳活動を予測することを行っている。

同定の方法では、低次視覚野の符号化モデル（生成モデル）を組み入れた判別モデルを用いることで、生成モデルを仮定しないSVMという判別モデルを用いた方法よりも、多クラス分類を高い正答率で実現した。一方で、同定の方法でも、対応表を作成するためには、つまりはこれが脳活動から刺激を推定するデコーダーであるわけだが、新規120の刺激セットが必要である。生成モデルの学習には使っていないデータなので新規刺激と呼んでいるが、デコーダーを構成する際には必要となる。つまり、既知の刺激セットから選択している点はクラス分類による方法と変わらない。

Kayらの研究が報告された同年の2008年に、宮脇、内田らにより、脳活動から 10×10 のパッチで表現される画像への回帰問題を解く再構成（Reconstruction）と呼ばれる方法でデコーディングした研究が報告された。視覚野における網膜座標配置の情報表現や、異なる受容野サイズによる情報表現など、脳科学で知られた符号化モデルを仮定し、脳活動から視覚刺激を予測する判別モデルを作成した。モデルを詳細に見ると、各パッチについての2値クラス分類を説明しているのだが、 2^{100} 通り以上の画像が表現できるモデルであり、連続空間への回帰問題を解いていると言っても良いだろう。クラス分類や同定では刺激（クラス）セットの情報が必要であったが、再構成では不要である。また、 10×10 のパッチで表現できる範囲内において任意の画像を表現する能力があり、少なくとも再構成できた画像程度の情報が脳活動に存在することをいえるだろう。

2. 脳活動解析の目的

本研究では、脳活動から知覚した顔を再構成し、再構成した情報を基に行動を予測することを目指した。ここでは、顔を知覚している時点での脳活動から、顔識別行動を予測するのであって、顔識別を行っている時の脳活動は使わない。本論文では、紡錘状回顔領域における顔知覚中の脳活動を用いて、いくつかの手法で顔の情報が取り出せるかを調べた。

3. fMRI 実験

3.1 fMRI 計測設定

fMRI 計測設定は、顔領域同定課題、顔識別課題ともに同様であり、次の通りである。3.0-Tesla Siemens MAGNETOM Trio A Tim scanner; TR, 3000 ms; TE, 30 ms; number of slices, 50 slices; Voxel size, 3 mm x 3 mm x 3 mm. fMRI 内で実験参加者が見る画像サイズは視野角にしておよそ 6.4 deg である。解析は SPM5 で行った。本研究は国際電気通信基礎技術研究所 倫理委員会の承認を受けた。

3.2 顔領域同定課題

まず、脳の高次視覚野を同定するための fMRI 実験を、顔領域同定課題を用いて行った。課題では画像を表示時間 1000 ms、時間間隔 500 ms で順次表示した。画像表示間にある 500 ms では黒の背景のみを表示した。被験者は順次表示される画像を注意してみつつ、1つ前の画像と等しい画像が表示されたらボタン押しにより回答する。表示した画像は、顔画像、物体画像、家画像、モザイク画像の4種である。1種類の異なる12画像が順次表示され、次に12 sec 間の黒背景表示となる(1ブロック)。黒背景表示の後に、異なる種類の画像のブロックが開始する。4種の画像が4ブロックずつランダムな順番で表示される。

以上の実験において、まず、有意にモザイク > ベースライン (*uncorrected*, $p < 0.1$) となった低次視覚領域を求めた。次に上で求めた低次視覚領域以外の脳領域から有意に顔 > ベースライン (*Family-wise error rate (FWE)*, $p < 0.05$) となる脳領域を選択した (クラスターサイズ > 50 voxels)。その結果、下側頭葉の 170 voxels が選択され、以降の再構成解析ではこの脳領域における脳活動を対象とした。解剖学的に見ても、第1次視覚野などの低次視覚野が含まれず、下側頭葉の領域だけが同定された。

表 10.1: 各手法における顔座標推定誤差

Method	Only regress	PCA+regress	ARD+regress	CCA +regress
Morph Step	19.9	5.33	8.02	4.88

その上で、第 8 章の 1. 顔識別課題で述べた課題を fMRI 計測中に行い、顔領域同定課題によって同定された領域の脳活動を以下の解析対象とした。

4. 顔の距離に相関する脳活動

図 8.7 において、同じ列にある顔画像は、それぞれの段の左端にある顔からの距離が等しくなるように設定している。ここでは、左端にある顔からの距離を回帰子として回帰分析をおこない、顔知覚モデルの距離に関連する部分を求めた。回帰分析をおこなう脳の範囲は、さきほど同定した顔領域である。その結果、距離に関連する部分が見つかった (図 10.2)。

5. 顔知覚座標の推定

189 試行のうち、188 試行を学習データ、1 試行をテストデータとした Leave one out で、脳活動が表現する顔知覚座標を推定した。推定には、4 つの方法を用いて推定した。1 つ目は、170 の脳活動から 38 次元を単に回帰分析する方法。2 つ目は、脳活動に PCA をしたあとに、回帰をする方法。3 つ目は関連度自動決定 (Automatic Relevance Determination; ARD) により選択された脳活動から回帰する方法。最後が正準相関分析 (Canonical Correlate Analysis; CCA) による方法である。本研究では、顔知覚は提示した顔画像と一致する実験であるので、推定した顔知覚座標と提示した顔の座標との差をとり、Morphing Step に換算して評価を行った。

その結果、CCA を用いた方法が最も推定誤差が小さく、平均 4.88 Step であった (表 10.1)

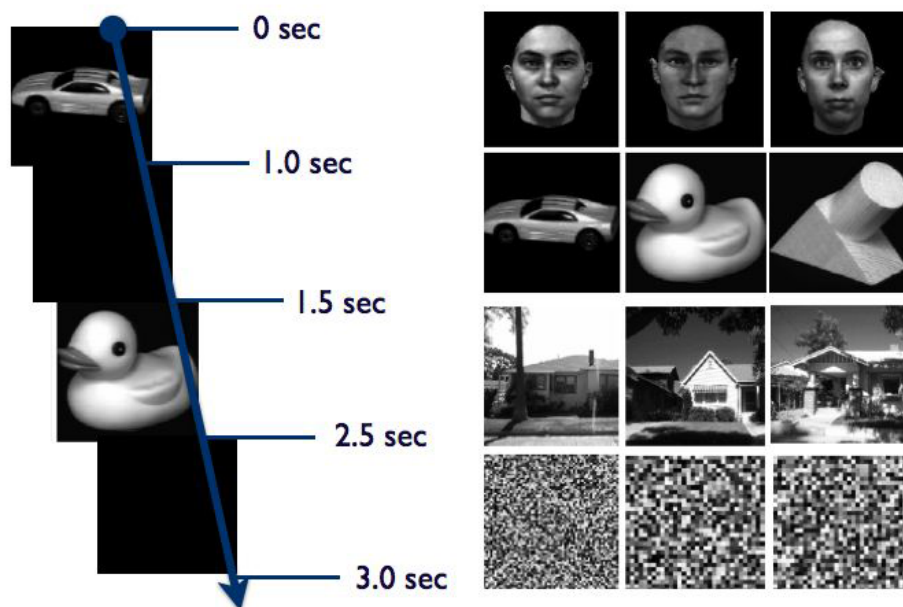


図 10.1: 機能的顔領野同定課題. 再構成で用いる顔選択的に活動する領野を同定するための実験を行った. 実験課題では画像を表示時間 1000 ms, 時間間隔 500 ms で順次表示した (左図). 画像表示間にある 500 ms では黒の背景のみを表示した. 被験者は順次表示される画像を注意して見ながら, 1 つ前の画像と等しい画像が表示されたらボタン押しにより回答する. 表示した画像は, 顔画像 (右図最上段), 物体画像 (右図上から 2 段目), 家画像 (右図上から 3 段目), モザイク画像 (右図最下段) の 4 種である (ただし, 顔領野を同定するのに用いたのは顔画像, モザイク画像を提示したときの脳活動のみである). 1 種類の異なる 12 画像が順次表示され, 次に 12 sec 間の黒背景表示となる (1 ブロック). 黒背景表示の後に, 異なる種類の画像のブロックが開始する. 4 種の画像が 4 ブロックずつランダムな順番で表示した.

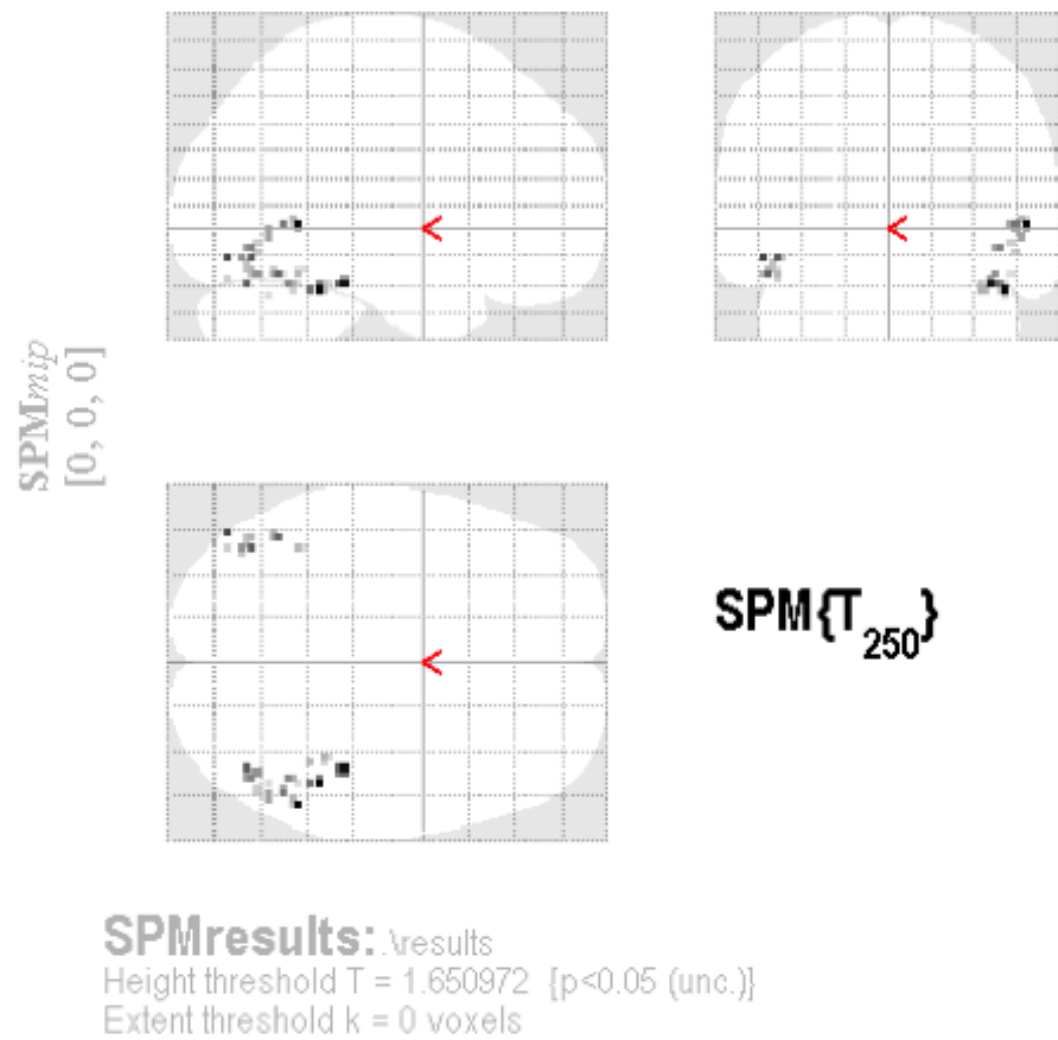


図 10.2: 顔の距離に相関する脳部位. 同定された顔領域内において, 顔の距離に相関して活動する部位である.

第11章 顔知覚についての議論

脳の高次視覚領野における顔知覚の情報表現を仮定した顔知覚モデルに基づいて作成した顔画像を用い、その顔画像を見せたときの顔識別行動が、顔識別行動モデルと良く一致することがわかった。これは、仮説とした顔知覚空間が脳でも実現されていることを示唆している。仮に、顔知覚浮かんが異なっていたとすれば、選択肢のちょうど中間にある顔で識別確率が0.5とならなかったり、左右非対称な識別確率になるはずである。

また、仮説としたモデルを検証するためには、識別行動だけではなく、さらに多くの作業仮説を検証する必要がある、今後の課題である。

また、脳活動の解析から、脳の顔領域で顔知覚空間の距離に相関する場所が見つかっており、顔知覚座標が脳活動から推定できることが示唆された。実際に、脳活動から顔知覚座標を推定してみたCCAを用いた方法では4.88 Step程度であった。ただし、詳細に調べてみると、推定された座標は顔知覚空間の原点周りに集中しており、良く予測できていない。今後の課題である。

宮脇、内田らの研究では、再構成できた画像を見せた画像と比較することで、再構成画像の評価を行った。一方で、本研究では脳活動から行動を予測をし、評価を行うことを考えている。

第3部

全体の議論

第12章 行動予測について

本論文では、2つの研究について述べたが、どちらも感覚刺激から行動の生成モデルをもちいて説明を行った。本章では、改めて、予測について議論を行いたい。

1. 予測の有用性について

1.1 生成モデル

理学研究において理解したいことは、現象の生成過程である。脳の情報処理で言えば、感覚刺激が最初に入力される領野での情報表現や、高次領野へと情報処理が進むに至る過程での情報の変換（アルゴリズム）を理解したい。

あるデータ y が得られたとして、パラメータ θ をもつパラメトリックな確率モデルで記述するとすれば、 $p(y|\theta)$ が生成モデル（Generative model）と呼ばれる。一方で、データからモデルのパラメータを推定する場合などに用いられる $p(\theta|y)$ は判別モデル（識別モデル, discriminative model）と呼ばれる。

ただし、個別のモデル全てについて完全に生成モデルと判別モデルのどちらかに別れるわけではない。たとえば、脳の情報処理について、宮脇、内田らの研究では、第1次視覚野の情報表現について生成モデルを考え、脳活動の情報から画像への再構成への判別モデルを用いて、画像予測を行ったと言える。

実験科学では、データを判別モデルによって説明し、得られた知見から帰納的推論をして生成過程の理解にたどりつく。基本的には判別モデルによる説明なので、数理モデルによる予測をおこなったからと言って、それだけでは生成過程の検証にはならない。であるから、理学研究では、生成過程の理解へつなげるために、単に判別モデルによって良い予測ができることを目指すのではなく、できるだけ仮説とする生成過程を組み込んだ判別モデルを作成し、予測精度を検証することが、生成モデルの検証に重要である。

現状では、脳の認知行動情報処理を調べる際に、演繹的推論を始めるための確実な知見は見あたらない。適当な知見から演繹的推論によって個別の高次情報処理について推論しても、他の研究者に新しい実験を促すような説得力を持たなさそうである。

演繹的推論のみによる方法をとるならば、すでに情報表現が理解されている領域における神経活動から目的の領域の神経活動を変換を求め、さらに神経活動から BOLD 信号への変換を求めて、情報表現を予測するという方法もある。

しかし、私たちは神経言語とでもいうべき、神経活動の情報表現原理を知らない。電気生理学では、発火率で情報が表現されているという前提に立つ研究は多いが、議論の余地は多く残されており、神経活動の情報表現原理についてはいまだに発展中の研究分野である。また、神経活動から BOLD 信号へのモデル化も試みられているが、まだ確定的ではない。研究対象である高次情報処理について調べるには、技術的な課題が多く残されている。

それでもなお、高次情報処理を研究対象と試みるためには、これまでの帰納的推論の方法を洗練させ、多次元情報である BOLD 信号のボクセルパターンデータと視覚などの多次元情報表現の関係を、効率よく帰納的推論を行う必要がある。

fMRI デコーディング研究において、Kay らの同定や宮脇、内田らの再構成、パラメトリックな判別モデルを考え、機械学習によってデータからモデルパラメータを帰納的に推論したと言える。また、宮脇、内田らの研究ではあらかじめ用意した基底表現によって視覚刺激画像が再構成されていたが、藤原らは基底表現をデータから推定する方法を発表した。このように、脳科学データに機械学習など情報科学の方法を有効に適用することで、新しい発見があると期待される。

また、理学研究では、生成過程の理解へつなげるために、単に判別モデルによって良い予測ができることを目指すのではなく、できるだけ仮説とする生成モデルを組み込んだ判別モデルを作成し、予測することが生成モデルの検証に重要である。

1.2 一貫性がないことへの弁明

非常に素朴な感覚からすると、私たちは日常的に現象を予測してみて、結果が何度となく一致した場合に、「分かった（理解した）」と感じる。一方で、情報量基準 AIC に一貫性（consistency）の無く、予測に最適化したモデルは真のモデルに一致しない。その点では、予測することは、真のモデルを知りたいとする科学の目的にそぐわないように考えられる。

第2部の系列予測の研究で、視標系列の次数同定について、一貫性（consistency）をもつ BIC と一貫性は持たないが良い予測をおこなうモデルを選択する赤池情報量基準 $AIC(p) = N \cdot \log \sigma_p^2 + 2(p+1)$ の2つを用いて比較した結果である（図 12.1）。

その結果、生成モデルの次数より高い次数においては AIC の値が BIC ほど大きくならなかった。このように、AR 過程のようにモデルが内包されている場合は、AIC は次数の高い（パラメータ数の多い）モデルを選択する傾向にあること

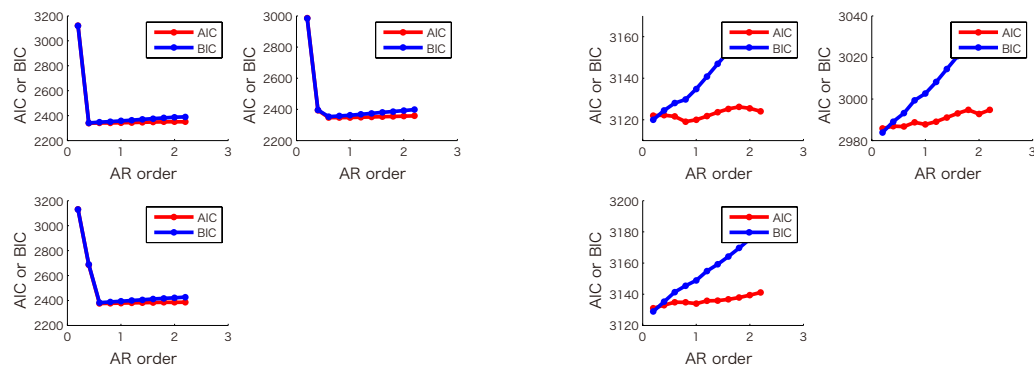


図 12.1: 視標系列の AIC, BIC による次数同定比較

本研究で用いた AR 系列と乱数系列について、横軸：AR 次数、縦軸：AIC もしくは BIC の値、とした結果。左にある 3 つの図が AR 系列の結果であり、右にある 3 つの図は乱数系列の結果である。AR 系列の 3 つの図のうち、左上が AR_1 系列、右上が AR_{21} 系列、左下が AR_{22} 系列である。乱数系列の 3 つの図のうち、左上が rnd_1 系列、右上が rnd_{21} 系列、左下が rnd_{22} 系列である。各図の赤線が AIC の値、青線が BIC の値である。

AR 系列については、全ての系列で生成モデルの次数が AIC, BIC の最小値となったが、乱数系列では、 rnd_1 系列について、生成モデル (0 次) より高次である 3 次の場合に AIC が最小となった。また、全体の傾向として、AIC の結果は生成モデルよりも高次である場合にも値は大きくなり、高次のモデルを同定する確率が高い。

が知られている。

ところで、一致性は、1. 真のモデルが階層的なモデル群の中に含まれている場合に、2. 観測データ無限大の極限で、選択したモデルが真のモデルと一致する確率が 1 になることを言う。

しかし、科学におけるモデル選択では、往往にして 1 の条件である階層的なモデル群からモデル選択をする場合ばかりではなく、また有限データからのモデル選択であり、一致性について厳密に議論できる問題の枠組みではない。また、AIC により高い次数が選ばれた場合もパラメータの値を分析すると高い次数のパラメータがゼロに近いなど、モデル自体の近さはそれほど変わらないので、評価できるのではないだろうか。

認知行動脳科学の情報処理モデルを考えたときに、精密な数理モデリングをできる知見はなく、概要的なモデリングの検証が行われている段階である。一致性がないことの偏りよりも、モデルが精密でないことの偏りが明らかに大きい。また、たとえ精密なモデルを立てたとしても、計測精度などの問題もあり検証できる実験は不可能である。

脳の情報処理について予測できる程度の表現、つまり数理的モデルによる表現を行い、予測に基づく評価を行うことが分野の発展に実効的ではないかと考える。

1.3 予測する手続きによる理解の確認

ヒトを理解するための構成論的方法を標榜し、ロボットを研究する研究者がいるが、彼らは「作ってみて気づくことがある」という言い方をする。確かに、経験的に、数理的にモデリングしてみたり、計算機でシミュレートするプログラムを作成する中で、パラメータの決定に必要な情報が足りていないことに気づくなど、作成することで必要な情報処理を理解することがある。自然言語で表現する場合には、正確に表現されていないために見落としていたり、十分に表現されているかを検証しにくかったりする。

科学全体としては、研究の論理や得られた知見が正しいことは、多くの研究者が議論することで保証されている。しかし、科学者個人が、行っている研究の論理や得られた知見の正しさを判断する場合はどうであろうか。私たちは、理解したと主観的に感じたとしても、実際には理解が十分ではなかったという場合もある。科学者もヒトであり、理解したと誤解してしまうことはあり得ることである。

その点で、予測するという手続きを取ることで、予測するのに必要な数理的表現での記述を求められるし、予測してみることで記述が正しいことを検証しやすくしていると言える。行動予測するためには、予測できる程度に数理的に記述する手続き、解析計算もしくは数値計算によって予測を行う手続き、また、予測と実際のデータを比較して評価する手続きが必要である。これらの手続きを実現することは、「モデルを定量的に予測できる程度の数理的記述で表現していること」を確認していることにもなると考えている。その点においても予測をすることの利点があると考え、微積分学における $\epsilon-\delta$ 論法のように、研究の論理や検証作業が正しく行われていることを確認しやすくしている方法論だと考えている。序論で述べたマーの3つの理解の水準と研究を照らしあわせると共に、行動予測を行うことで、その研究で得られる理解を計りながら研究を進めることができると考えている。

2. 「行動」の予測

脳活動を計測すれば、実験者が行動（運動）と脳活動の2つを観測として持つことが出来る。ところで、脳活動の観測で得られるのは、神経活動の発火率であったりfMRI信号であり、脳内で起こる物理的挙動であって、実験者が知りたい情報表現そのものではない。そのため、例えば、視覚刺激から脳活動の情報表現を予測したとしても、評価するための比較対象が無い。つまり、如何に脳活動計測が発展して精細な観測が得られたとしても、情報表現の予測だけを評価することはできない。

神経活動の発火や筋肉収縮など生体の完全な計測によって、行動に至までの物理的な現象が観測でき、物理的現象として行動を予測できたとして、脳を理解したことになるだろうか、おそらく、それだけでは理解したことにはならない。私たちが理解したいのは物理的な現象だけではなく、物理現象が表現する情報である。

知覚や認知の情報処理を対象とした従来の研究では、視覚刺激を与えるだけで、必ずしも行動を伴わない実験課題が多く行われている。行動が伴う実験課題である場合も、行動の観測はきちんと対象とする情報処理が行われていることを確認するための補足的な情報に過ぎないことが多い。この様な実験では、脳活動のおこる時間やある領野の脳活動が、仮説としたモデルとの相関することを調べることで、情報の有無を検証している。たとえば、1990年代から非侵襲の脳活動計測であるfMRI計測が発展し、ヒトの脳活動を良い空間解像度で観測できるようになった。2008年に宮脇、内田らは、10x10のパッチで表現される画像ならば、任意の視覚情報を予測（彼らは再構成（Reconstruction）と呼んだ）できるようし、fMRI計測から得られた脳活動から実験参加者の見ている画像を予測する方法を発表した。

しかし、その評価は、提示した視覚刺激画像と予測画像を画像として並べて表示する方法や提示した視覚刺激と比較する方法であった。もしくは、先行研究との比較を明にするために、多数の画像から実験参加者が見ている画像をどの程度推定できるかをその評価であるため、脳科学研究における評価としては不十分であった¹。

なぜならば、脳の情報表現は脳内での情報処理がされた結果であるから、すでに見せた視覚刺激の情報そのものではなく、予測画像を視覚刺激と比較することは妥当ではない。脳科学研究として予測したいのは脳活動で表現されている情報であり、脳活動がどのような情報を、どの程度に持っているかである。脳活動計測で観測できるのはfMRI計測で言えばBOLD信号であり、当然であるが、脳内で表現されている真の情報は観測できないので、予測した視覚刺激画像を直接に比較する対象はない。

認知行動の脳科学研究において、予測して評価することができるのは、行動（運動）だけである。つまり、知覚や認知の情報処理には必ずしも行動が伴うことはないが、予測に基づく検証を行うためには、行動を観測できる実験課題を考案する必要がある。これらの問題を解決するために、目的の領野において、観測として得られたBOLD信号から情報表現への変換を考え、情報表現から行動を予測した。また、予測した行動を観測された実験参加者の行動と比較することで、評価する。

¹宮脇、内田らの研究は、第1次視覚野など低次視覚野からの視覚情報予測を行っている。第1次視覚野は、大脳皮質では最初に視覚情報が入力される領野であり、網膜像の情報をほぼそのまま持っていると考えられている。

1990年代に入り、fMRI脳活動計測が出来るようになって認知行動の研究を行った多くは、心理学者であった。それは、差分解析法など、従来の心理学研究で使われてきた解析を援用することができたことも大きかった。fMRI計測の最大の利点は、電気生理学による方法と比較して、全脳を計測できることである。そのデータは、多次元である。今後さらに計測技術が向上する中で、取得した多次元データを分析することが脳の情報処理の解明につながるだろう。

本論文では、すでに発達や学習によって獲得している情報処理について述べてきたが、今後の発展として、時間発展を考慮した脳機能の記述が必要であると考えている。特定の機能を担う領野やニューロンの発見は重要ではあるし、ある程度の機能局在が存在する。しかし、ヒトの脳は常に環境と相互作用しながら変化しているという脳機能理解が必要であり、時間発展を考慮した脳機能の記述も必要である。

また、脳においてどの様に機能が実現されているかを考えると、ある機能がある領野で表現されている場合に、その機能は領野内における局所神経回路として実現されているように、機能によっては領野間結合の回路として実現している事もある。今後のfMRI研究では、単に機能局在を調べるのではなく、領野間結合を調べ、機能として表現することも必要であると感じている。

補足

全タスク・全実験参加者の予測系列

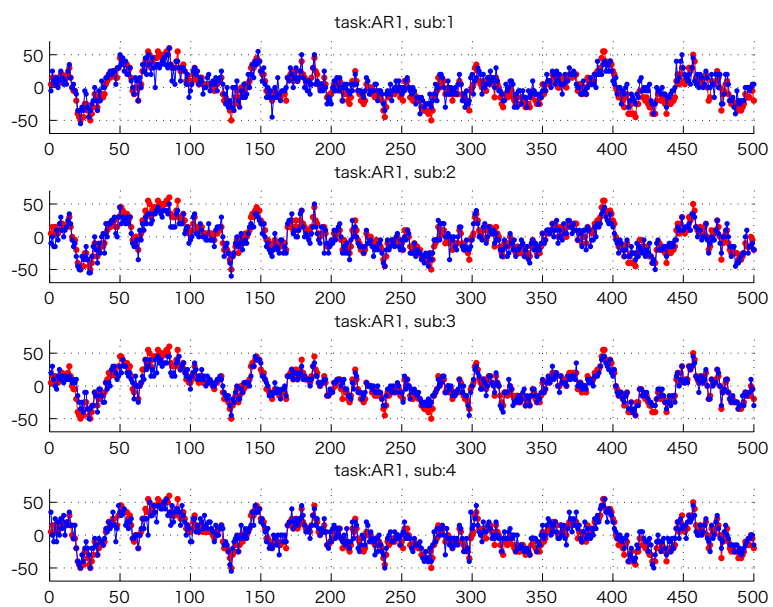


図 12.2: 実験参加者の予測系列・ AR_1 タスク

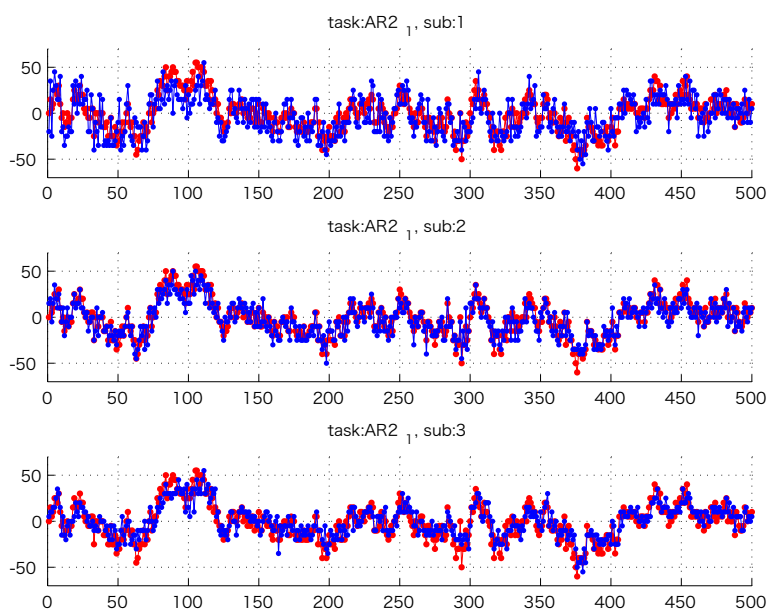


図 12.3: 実験参加者の予測系列・ AR_{21} タスク

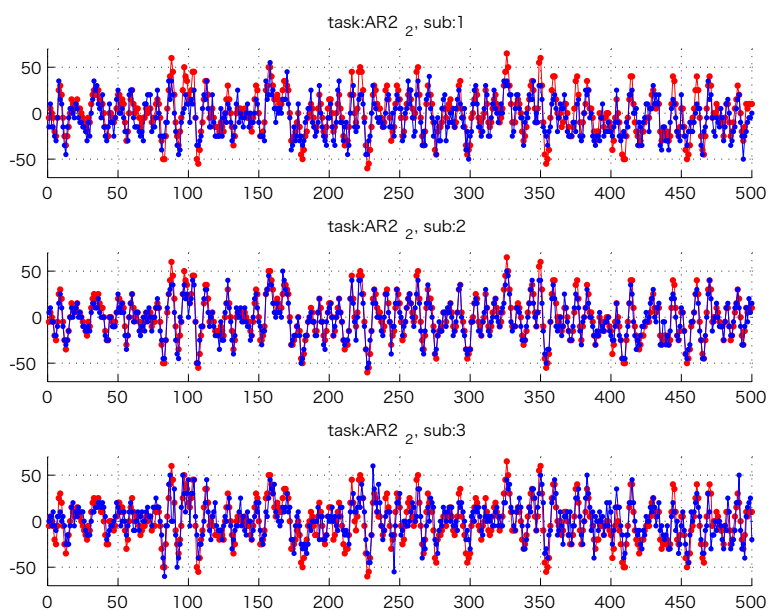


図 12.4: 実験参加者の予測系列・ AR_{22} タスク

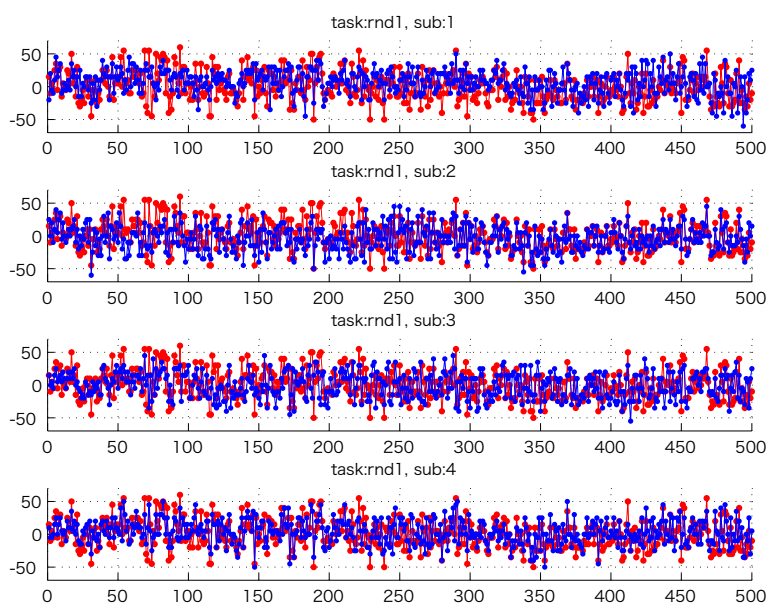


図 12.5: 実験参加者の予測系列・ rnd_1 タスク

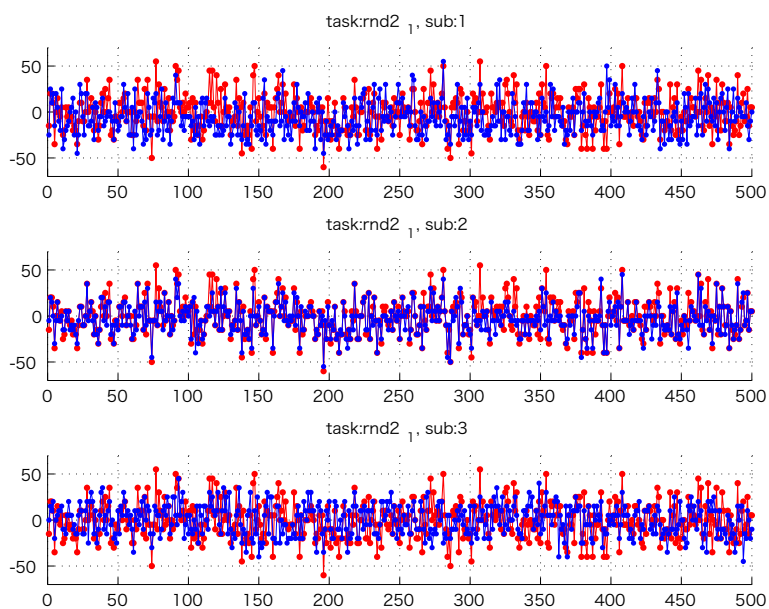


図 12.6: 実験参加者の予測系列・ rnd_1 タスク

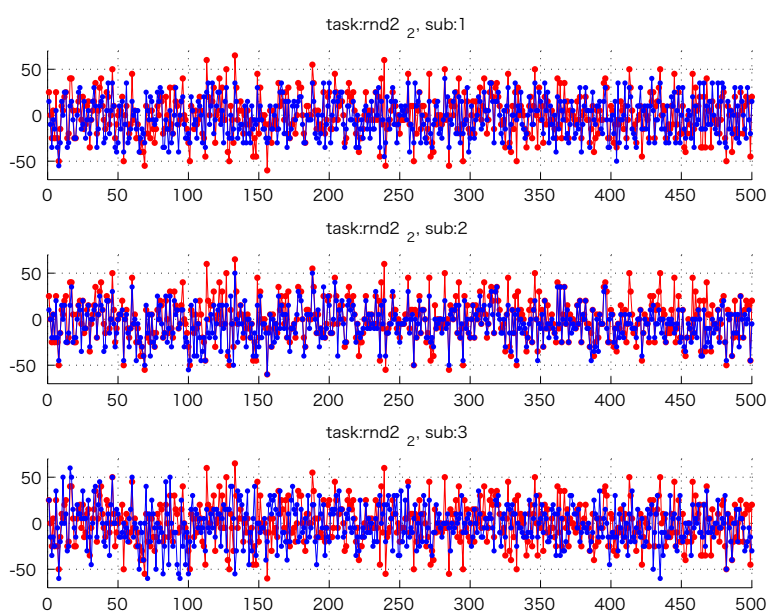


図 12.7: 実験参加者の予測系列・ rnd_{22} タスク

全タスク・全実験参加者の行動予測

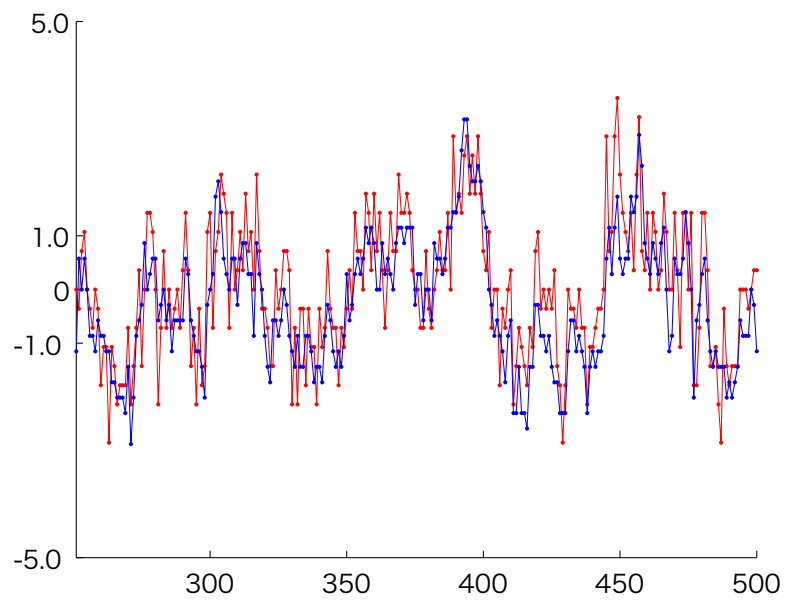


図 12.8: 実験参加者行動の予測・ AR_1 タスク, KT

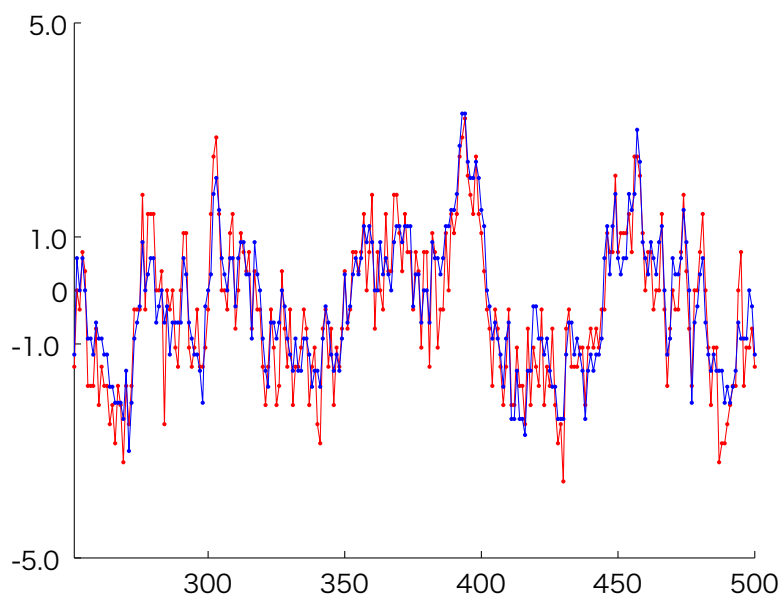


図 12.9: 実験参加者行動の予測・ AR_1 タスク, MO

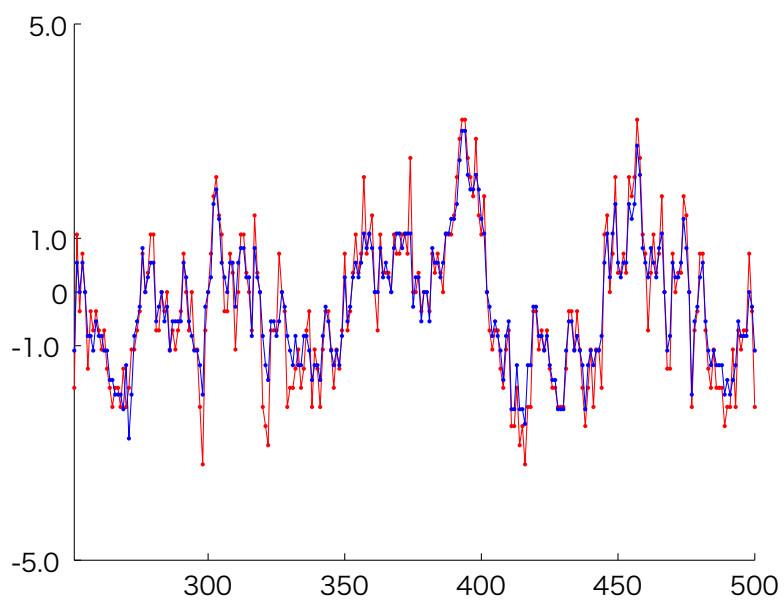


図 12.10: 実験参加者行動の予測・ AR_1 タスク, NN

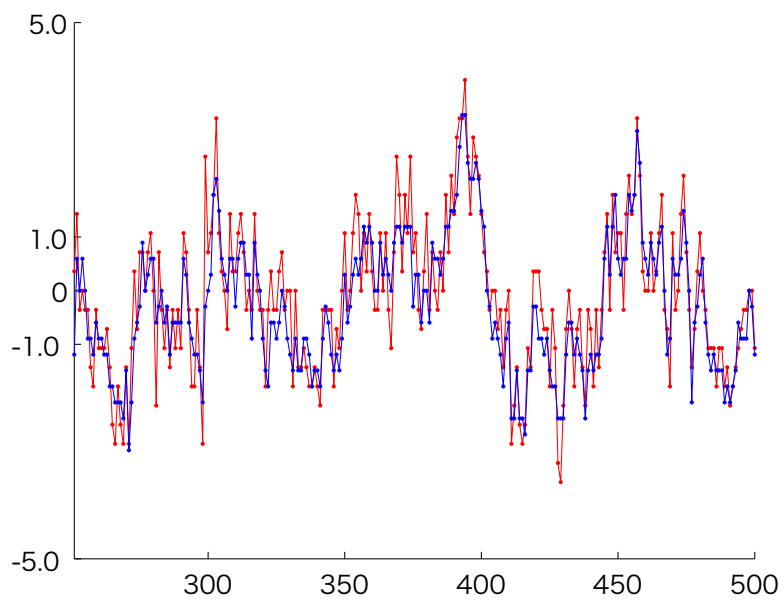


図 12.11: 実験参加者行動の予測・ AR_1 タスク, TY

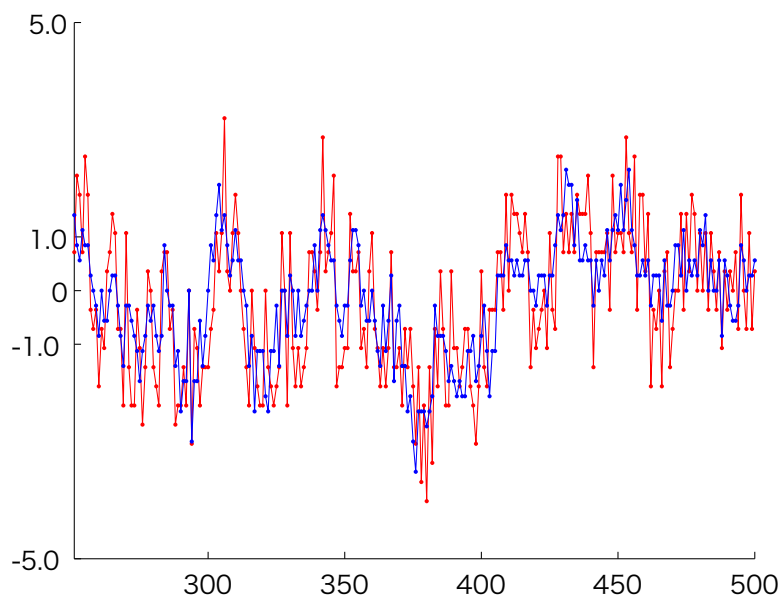


図 12.12: 実験参加者行動の予測・ AR_{21} タスク, KY

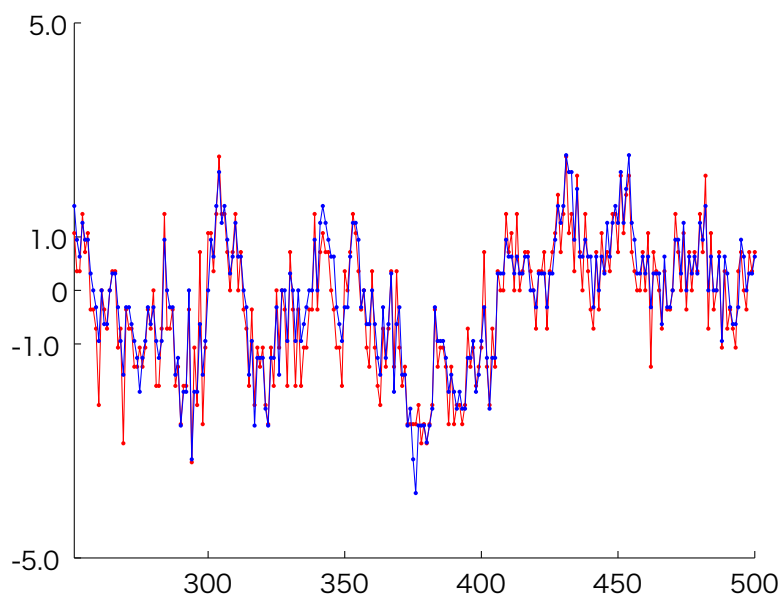


図 12.13: 実験参加者行動の予測・ AR_{21} タスク, MO

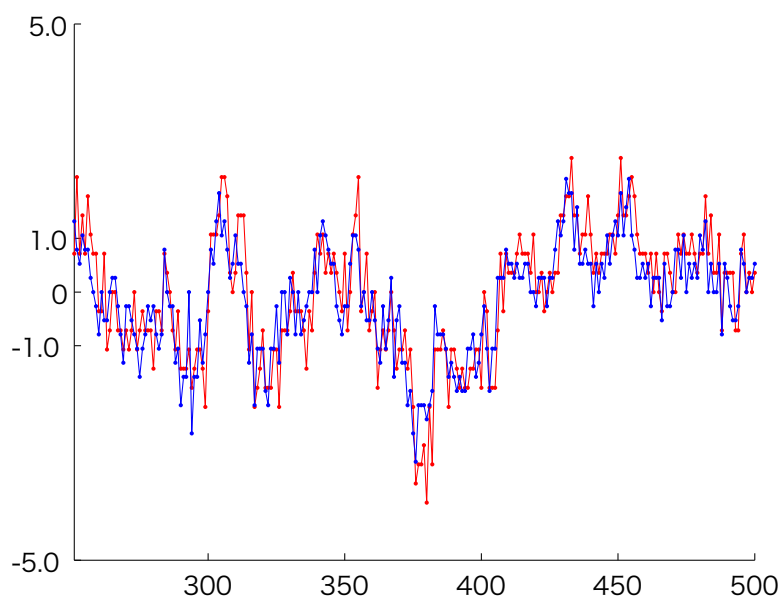


図 12.14: 実験参加者行動の予測・ AR_{21} タスク

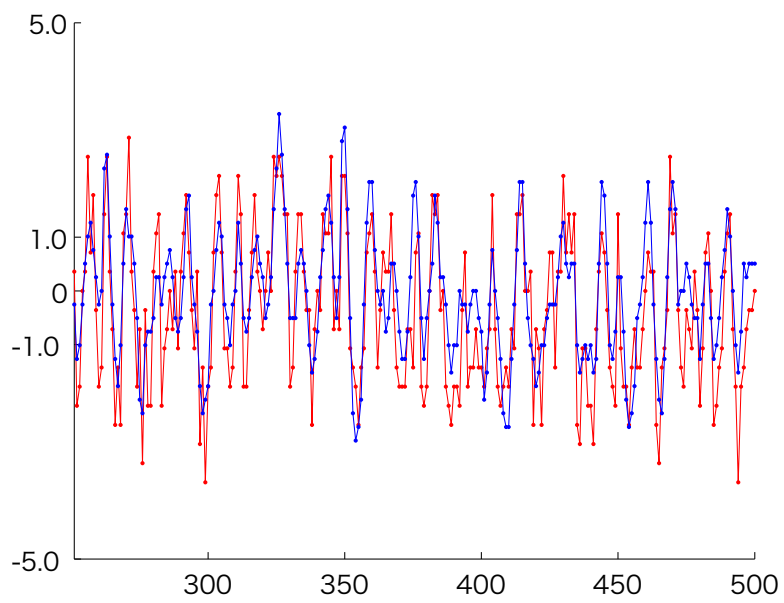


図 12.15: 実験参加者行動の予測・ AR_{22} タスク, KY

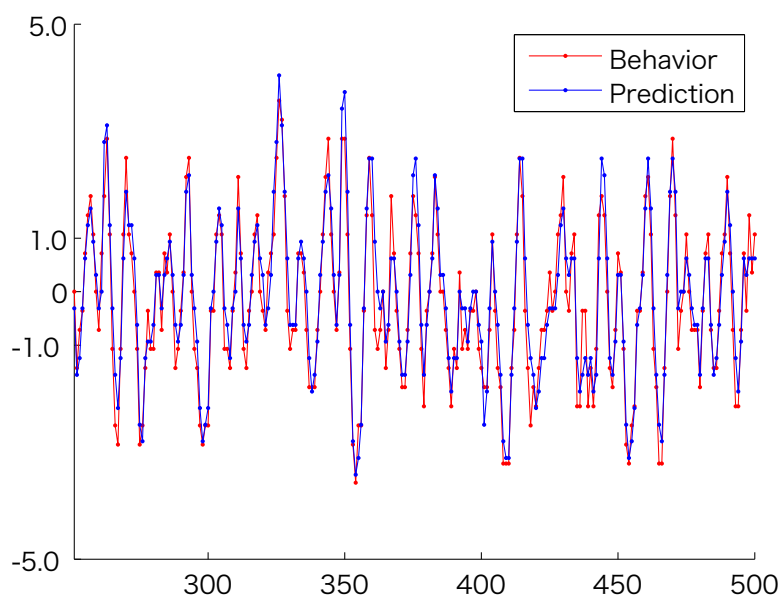


図 12.16: 実験参加者行動の予測・ AR_{22} タスク, MO

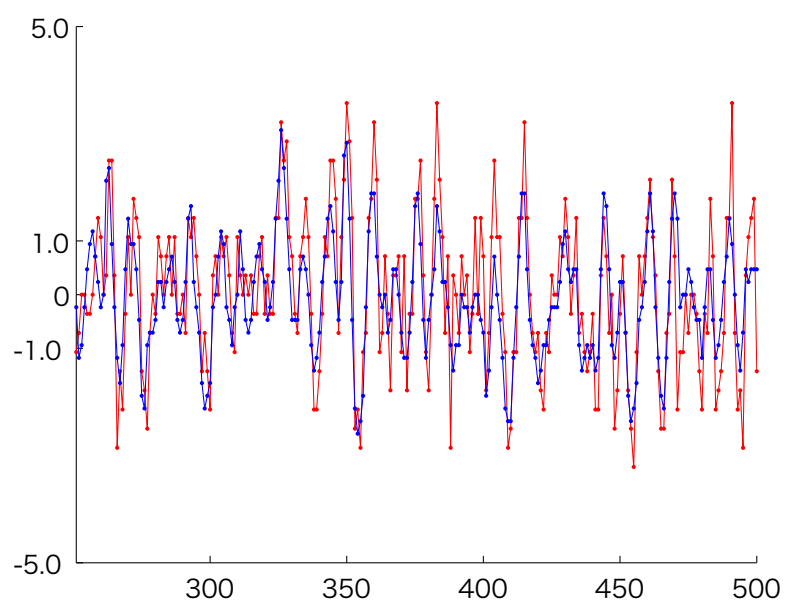


図 12.17: 実験参加者行動の予測・ AR_{22} タスク, NN

全実験参加者の顔識別行動

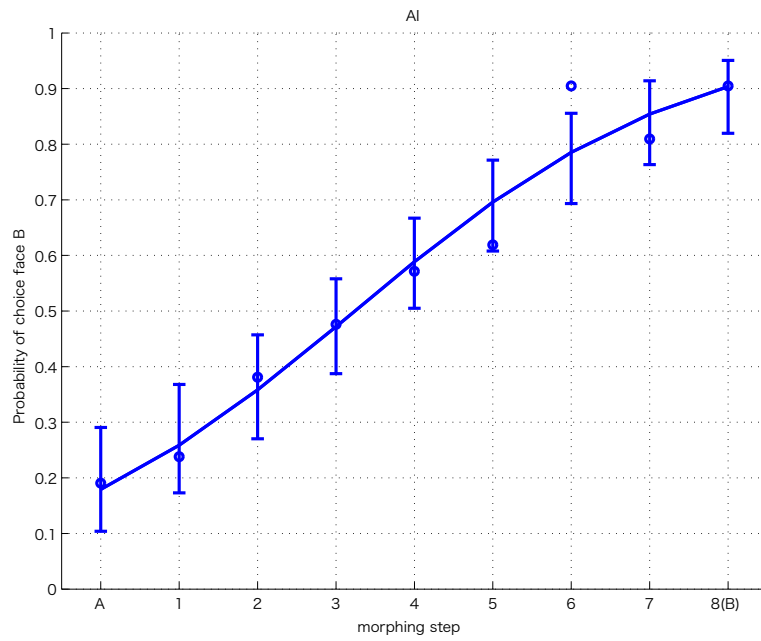


図 12.18: 顔識別行動, AI.

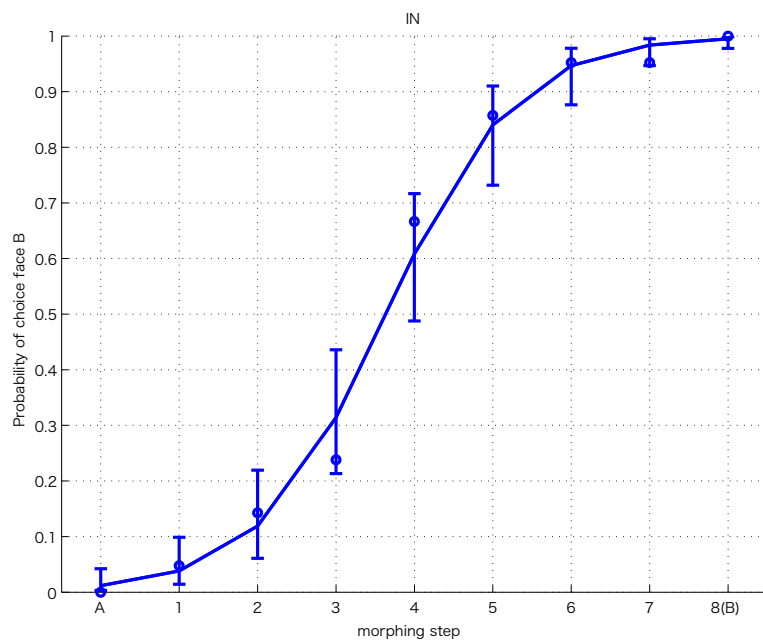


図 12.19: 顔識別行動, IN

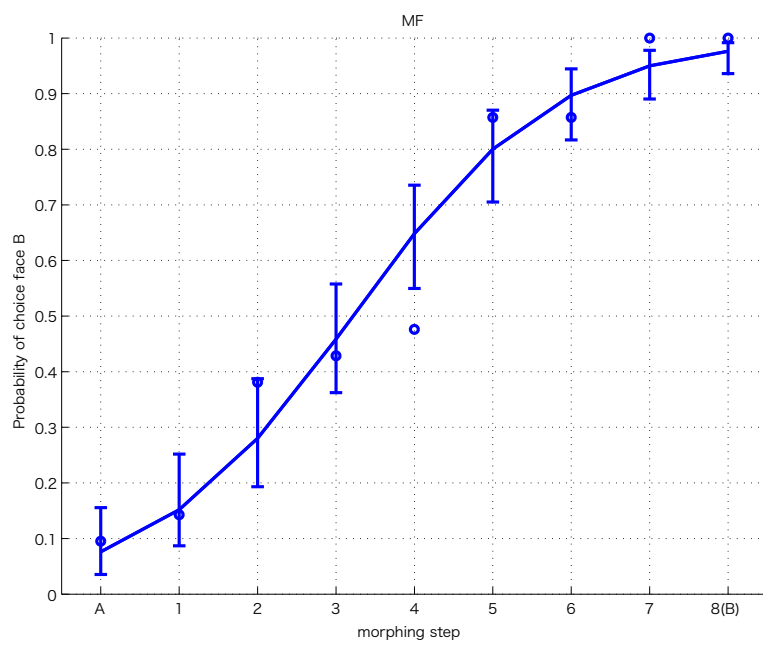


図 12.20: 顔識別行動, MF

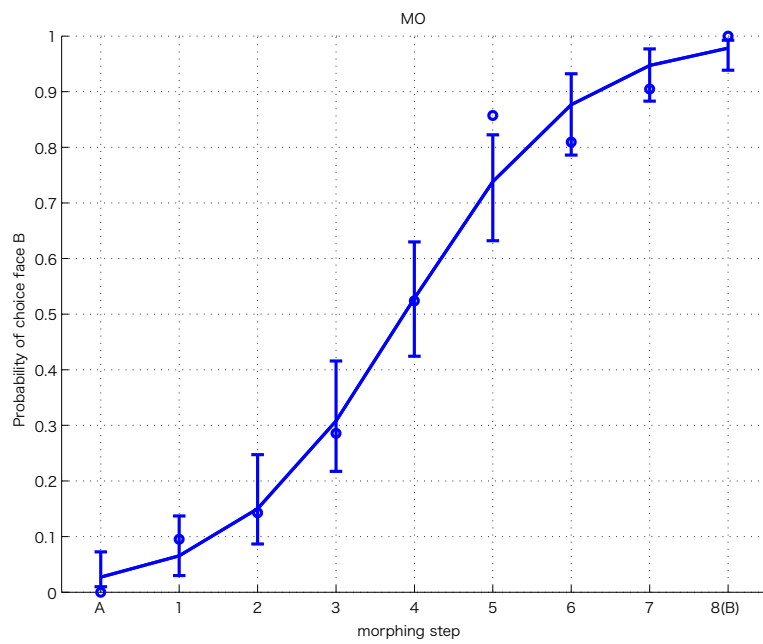


図 12.21: 顔識別行動, MO

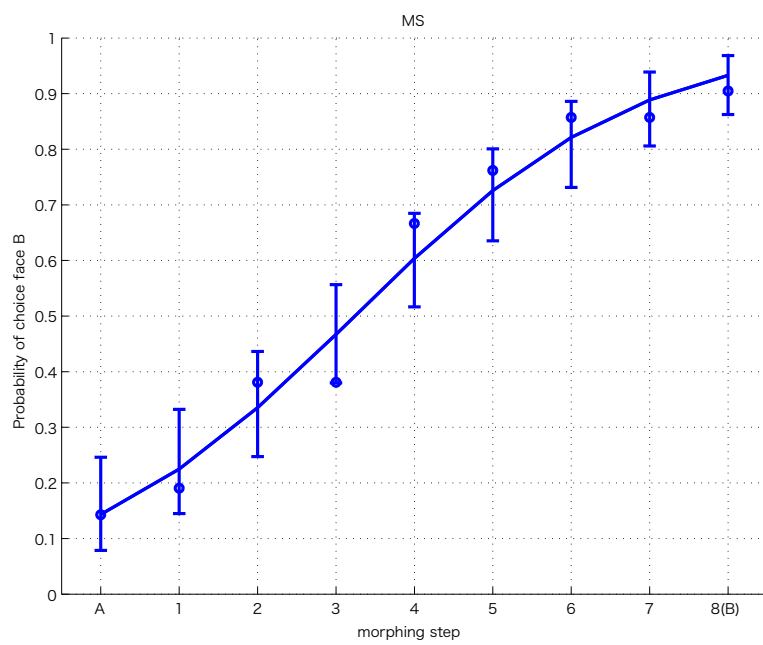


図 12.22: 顔識別行動, MS

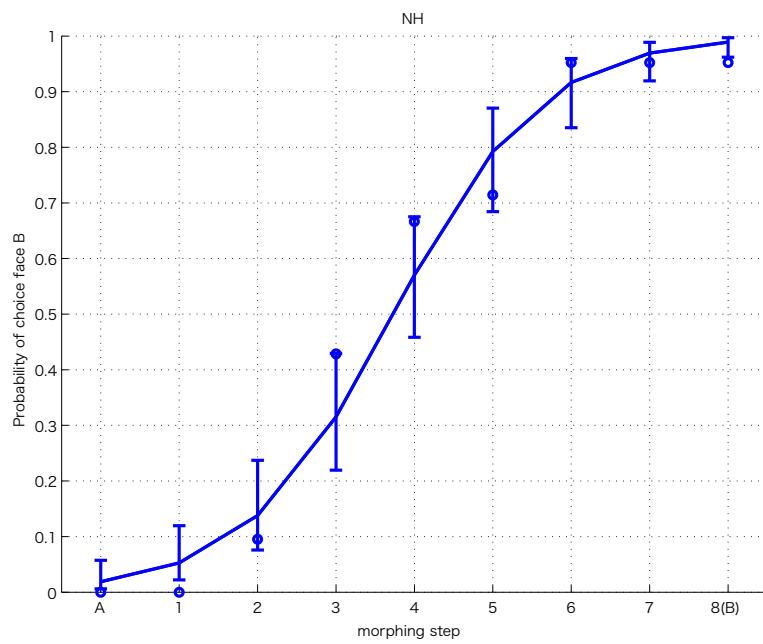


図 12.23: 顔識別行動, NH

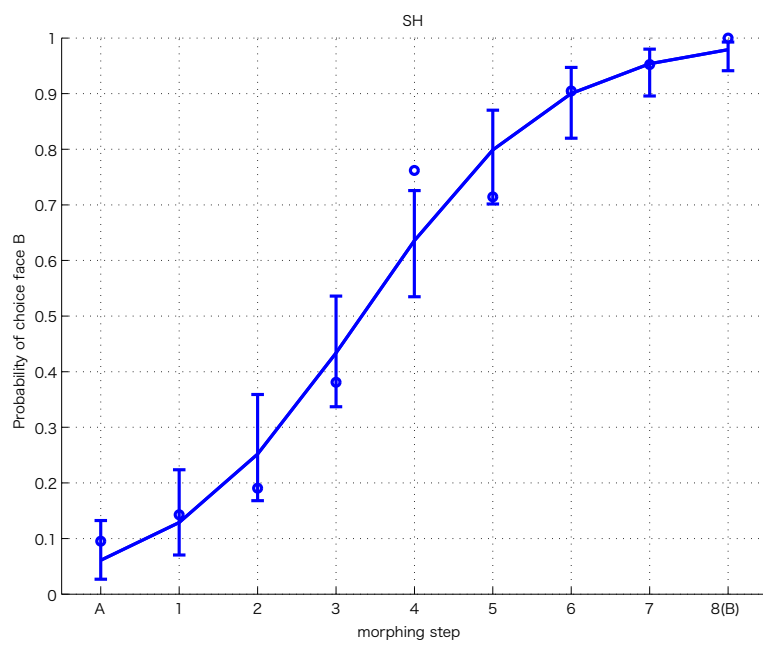


図 12.24: 顔識別行動, SH

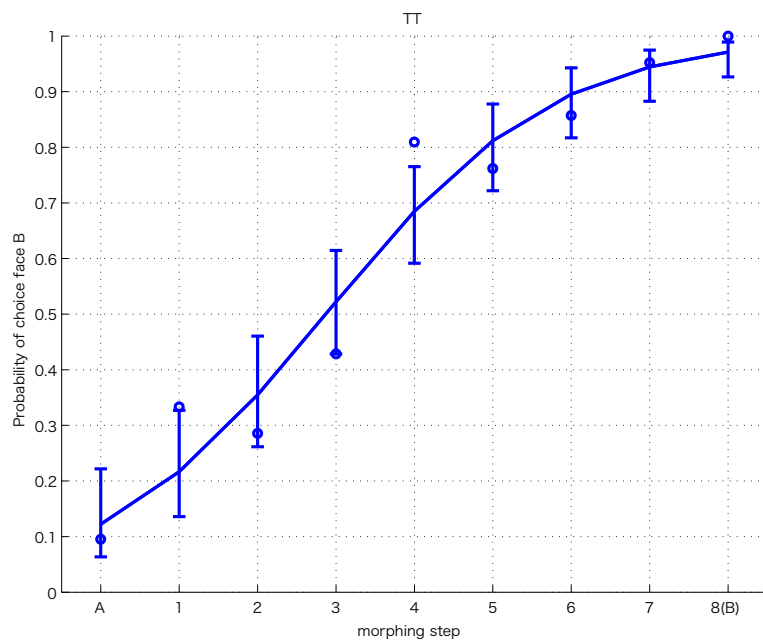


図 12.25: 顔識別行動, TT

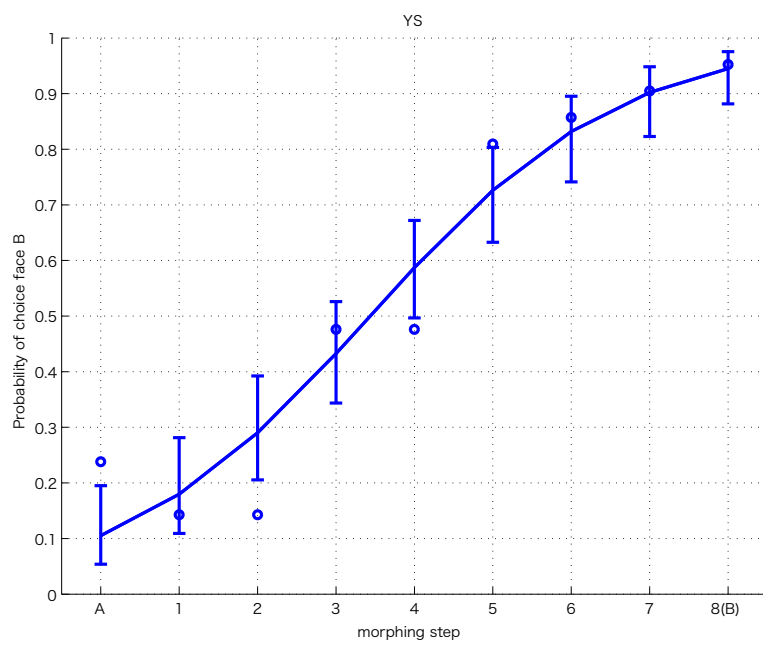


図 12.26: 顔識別行動, YS

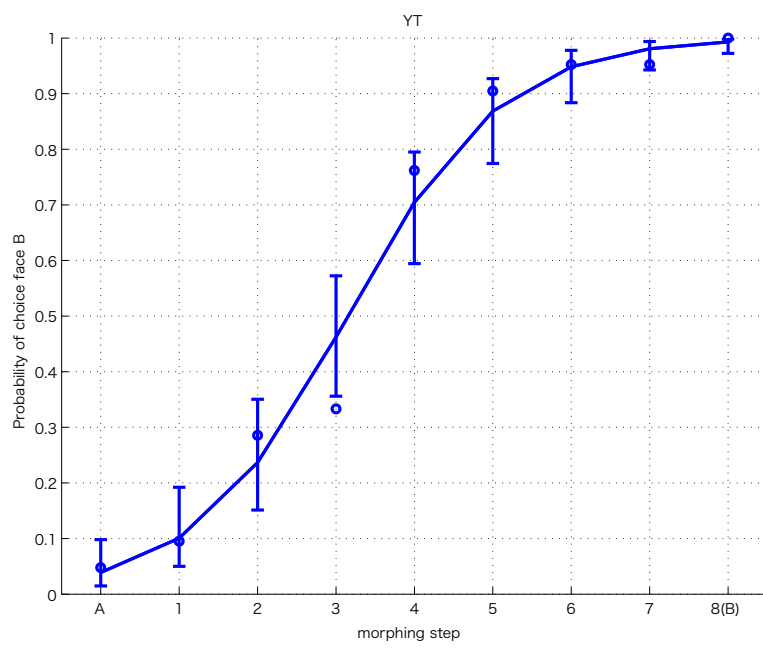


図 12.27: 顔識別行動, YT

謝辞

石井 信 教授，湊 小太郎 教授，柴田智広 准教授には私の学位審査をしていただきました，ありがとうございました。

以下，普段お呼びしている敬称で失礼させていただきます。

石井先生，柴田さんには，学生である間，また私が京都大学に移動した後も事ある毎にご指導いただきました。この博士論文が完成するまでに多大なご助力をいただいたことを，ここに感謝いたします。私は博士後期課程から研究分野を変えて論理生命学講座に進学してきましたが，柴田さんには修士の頃からご指導いただいていたいました。

私が3年間を過ごした論理生命学講座では，深夜になっても部屋の電気が消えることがなく，いつでも誰かが研究をしていました。また，たくさんの方が研究していたので，様々な仕事の進め方を目の辺りすることが出来ました。ここでの経験は今後仕事を進めていく上での原点になると思います。また，論理生命学講座では，数理統計，機械学習，人工知能，ロボット，画像情報処理，システム生物学，認知科学など普通では異なる研究室で行われている研究が1つの研究室で行われていました。ですが，これらは全て，ヒトを研究する上での知識と方法論について重要な手がかりを与えてくれます。

また，論理生命学講座のみなさんとは公私にわたりおつきあいしていただきました。個人的にも交流して励ましていただいたことは，普段の生活と研究が隣り合っている私にとっては，まわりまわってこの論文を完成するための大きな助けとなりました。論理生命学講座を離れた後も交流が続いている友人が多く嬉しく思っています。以下，論理生命学講座を通じて出会い，助けていただいた方々に感謝を述べたいと思います。

谷本さん，那谷さん，藤澤さん，齋藤さん，林さん，中川さん：石井先生の秘書の方々です。研究に関わる様々な手続きのお手伝いをいただきました。特に，私は実験などに関するいくつかの書類提出が必要でした。時には誤って提出した書類も知らぬところで修正していただいたようで，研究に専念できる環境を作って

いただきました。また藤澤さんには、学術論文の英語の校正でも助けていただきました。研究が研究者だけでは成立しないことを強く意識する機会を与えてもらいました。

作村さん：今回の学位審査の手続きをはじめ、私が気づかないところで助けてもらったと想像しています。

大羽さん、竹之内さん：論生命講座の1年目は居室が一緒だったので夜な夜な議論をしたこともありました。数理統計でわからないときも実効的なご助言をいただきました。学生の際は、勝手ながら、研究上のよりどころにしていたように思います。

前田さん：非常に素朴な質問をしてもらって、答えられないことがしばしばありました。その度に改めて研究の土台を見直す機会になりました。

五十嵐さん：勉強会や夕方から行っていたバスケットなどなど、こちらが少し気の乗らない顔をした時にも、しつこく誘ってもらいました。最初は気分が乗らなくても、やってみたら良かったということは頻繁にある事です。私が論理生命にいた最後の2年間、五十嵐さんは卒業されていなかったのですが、五十嵐さんの影響が研究室のあちらこちらに残っていたように思います。私にとって居心地の良い研究室にしてくれました。

平山さん、坂東さん：本論文の後半部の研究にも関わっていただきました。公私ともにお世話になったおふたりです。平山さんとは良くお酒を飲みに行きましたし、坂東さんとはソフトボールなどで一緒に遊びました。

兼村くん、植野くん：2人が京都大学に移るまでは、論理生命学講座メンバー中で最も長い時間を一緒に過ごしたと思います。研究上のアイデアにもなっていないアイデアをいろいろと聞いてもらいました。他の人には恥ずかしくて聞けない数理の初歩的な質問ができるのもこの2人でした。大学院生活で気の置けない友人が出来ました。改めて敬称をつけて呼ぶのは照れくさいのですが、感謝を示したいと思います。ありがとうございました。

本田くん：論理生命学講座内では実質的に彼だけが同じ学年で、強く意識していたと思います。ホワイトボードの前で議論しながら研究を進める方法は刺激になりました。敬称をつけて呼ぶのは変な気分です。

塚田さん：粘り強く、ですがそれを表に出さず飄々と明るく研究されていたことが印象に残っています。公私ともにお世話になりました。

吉田さん、森本くん、大塩さん：fMRI計測が必要な研究を始め、1人で右往左往していた頃にいろいろと助けてもらいました。

為井くん：論理生命学講座に入ったとき、最初に声をかけてもらったと思います。いろいろな場面でさりげなく現れて助けてもらいました。今回の学位論文の手続きでも助けてもらいました。サッカーを通じての交流もありました。

南部くん：今回の学位審査に関わる手続きでも、私が京都大学にいたので学内か

らしか得られない資料を送ってもらったりと気にかけていただき、大きな助けになりました。公聴会にも来てもらいました。為井くん同様、サッカーを一緒にやりました。

大倉くん、吉田くん：公私で言えば、「私」での交流が多かった2人です。おかげで研究にも弾みがつきました。

井原さん：他人の評価にぶれることなく研究している姿勢に影響を受けました。公聴会にも来てもらいました。

中村さん、金子さん、川脇さん、藤原くん、大塚くん、大佐賀さん、福永くん、孫野くん、殿谷くん、牧野くん、丸山くん、竹田くん、南條くん、山下くん、行縄さん、藤田さん、Allan、人見くん、矢野さん、深田くん、鈴木さん：ありがとうございました。

脳科学の生理学的や心理学的知見や実験技術に関しては非常に未熟だった私にとっては、国内外の会議で出会った学外の友人との議論で教わることも多くありました。出会った頃の所属を併記します。

佐々木くん（順天堂大学）：私が脳科学を始めた最初の年に出会って以来、事ある毎に研究会などで一緒になり、相手をしてもらいました。

石川さん（大阪大学）、番くん（京都大学）：関西システム神経科学若手の会で出会い、現在でも気軽に研究やその周辺のことについて話せる友人です。

野村さん、金津さん、中園くん（以上、京都大学）：京都大学で研究科を横断して集まって行っている論文読み会に参加してもらっています。脳科学は、学際的分野で非常に広範な分野にわたって研究されています。多くのことを知った上で研究を進めたい私にとっては、この論文読み会は良い機会になっています。

梅田さん、土井くん、稲垣さん（以上、大阪大学）柴田くん（ATR・NAIST）：関西システム神経科学若手の会で一緒でした。

本論文の研究を完成させる後半の2年間は、京都大学ナノメディシン融合教育ユニットでの仕事をしながらとなりました。京都大学 医学研究科 認知行動脳科学分野のみなさんや特任教員にお世話になりました。

河野先生、竹下さん：はじめての仕事をいろいろと助けてもらいました。

三浦さん：眼球運動の心理物理実験に参加させていただき、改めて実験手技を教わりました。ナノメディシン融合教育ユニットの仕事でも様々にサポートしていただきました。また、また、研究室で一緒になる前から研究会などでの私の発表

を聞きに来ていただいていた、多大に評価していただくこともあり、励みになりました。

小川さん：論文読み会を通じて、生理学的な考え方、電気生理学の論文の読み方などを教わりました。

松田さん、宮野さん、桑島さん、原田さん、外波さん：ナノメディシン融合教育ユニットの特任教員として、中心的な役割を果たしておられ、いろいろと助けてもらいました。松田さんとは、同じコースを担当ということもあり、カリキュラムの作成など一緒に作業することも多かったです。仕事を始めて間もない私をいろいろと助けてもらいました。宮野さんは、仕事の進め方や考え方などでも共感できる部分が多く参考になりました。また、宮野さんを通じてたくさんの出会いに恵まれたことも嬉しかったことです。

最後に、妻である鹿内友美に感謝します。論理生命学講座で出会い、学生であった頃から今も支えてもらっています。

参考文献

- [1] H. Akaike. On entropy maximization principle. *Applications of Statistics*, 27(41), 1977.
- [2] A.T. Bahill and J.D. McDonald. Model emulates human smooth pursuit system producing zero-latency target tracking. *Biol Cybern*, 48:213–222, 1983.
- [3] G. R. Barnes and P. T. Asselman. The mechanism of prediction in human smooth pursuit eye movements. *J Physiol*, 439:439–461, 1991.
- [4] Olaf Blanke and Margitta Seeck. Direction of saccadic and smooth eye movements induced by electrical stimulation of the human frontal eye field: effect of orbital position. *Experimental brain research Experimentelle Hirnforschung Expérimentation cérébrale*, 150(2):174–83, May 2003.
- [5] K H Britten, M N Shadlen, W T Newsome, and J A Movshon. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci*.
- [6] C J Bruce and M E Goldberg. Primate frontal eye fields. i. single neurons discharging before saccades. *Journal of Neurophysiology*, 53(3):603–35, Mar 1985.
- [7] V Bruce and A Young. Understanding face recognition. *British journal of psychology*, Jan 1986.
- [8] Mark J Buckley, Farshad A Mansouri, Hassan Hoda, Majid Mahboubi, Philip G F Browning, Sze C Kwok, Adam Phillips, and Keiji Tanaka. Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. *Science*, 325(5936):52–8, Jul 2009.
- [9] M. J. Catalan, M. Honda, R. A. Weeks, L. G. Cohen, and M. Hallett. The functional neuroanatomy of simple and complex sequential finger movements: a pet study. *Brain*, 121 (Pt 2):253–264, 1998.

- [10] L L Chen and S P Wise. Neuronal activity in the supplementary eye field during acquisition of conditional oculomotor associations. *Journal of Neurophysiology*, 73(3):1101–21, Mar 1995.
- [11] L L Chen and S P Wise. Supplementary eye field contrasted with the frontal eye field during acquisition of conditional oculomotor associations. *Journal of Neurophysiology*, 73(3):1122–34, Mar 1995.
- [12] P. Dallos and R. Jones. Learning behaviour of the eye fixation control system. *IEEE Trans Autom Contr*, AC-8:218–227, 1963.
- [13] M C Dorris and D P Munoz. Saccadic probability influences motor preparation signals and time to saccadic initiation. *J Neurosci*, 18(17):7015–26, Sep 1998.
- [14] M C Dorris, M Paré, and D P Munoz. Neuronal activity in monkey superior colliculus related to the initiation of saccadic eye movements. *J Neurosci*, 17(21):8566–79, Nov 1997.
- [15] J.C. Eliassen, T. Souza, and J.N. Sanes. Human brain activation accompanying explicitly directed movement sequence learning. *Exp Brain Res*, 141:269–280, 2001.
- [16] P T Fox and M E Raichle. Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. *Proc Natl Acad Sci USA*, 83(4):1140–4, Feb 1986.
- [17] Winrich A Freiwald, Doris Y Tsao, and Margaret S Livingstone. A face feature space in the macaque temporal lobe. *Nat Neurosci*, 12(9):1187–96, Sep 2009.
- [18] K J Friston. Bayesian estimation of dynamical systems: an application to fmri. *NeuroImage*, 16(2):513–30, Jun 2002.
- [19] B Gaymard, S Rivaud, J F Cassarini, T Dubard, G Rancurel, Y Agid, and C Pierrot-Deseilligny. Effects of anterior cingulate cortex lesions on ocular saccades in humans. *Experimental brain research Experimentelle Hirnforschung Expérimentation cérébrale*, 120(2):173–83, May 1998.
- [20] P Van Gelder, S Lebedev, P M Liu, and W H Tsui. Anticipatory saccades in smooth pursuit: task effects and pursuit vector after saccades. *Vision Res*, 35(5):667–78, Mar 1995.

- [21] J A Van Gisbergen, A J Van Opstal, and A A Tax. Collicular ensemble coding of saccades based on vector summation. *Neuroscience*, 21(2):541–55, May 1987.
- [22] P W Glimcher and D L Sparks. Movement selection in advance of action in the superior colliculus. *Nature*, 355(6360):542–5, Feb 1992.
- [23] M E Goldberg and M A Segraves. The visual and frontal cortices. *Rev Oculomot Res*, 3:283–313, Jan 1989.
- [24] Kalanit Grill-Spector, Nicholas Knouf, and Nancy Kanwisher. The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci*, 7(5):555–562, May 2004.
- [25] P. Haggard and J. Richardson. Spatial patterns in the control of human arm movement. *J Exp Psychol Hum Percept Perform*, 22(1):42–62, 1996.
- [26] P J Hancock, A M Burton, and V Bruce. Face processing: human perception and principal components analysis. *Mem Cognit*, 24(1):21–40, Jan 1996.
- [27] D P Hanes and R H Wurtz. Interaction of the frontal eye field and superior colliculus for saccade generation. *Journal of Neurophysiology*, 85(2):804–15, Feb 2001.
- [28] J Haxby, E Hoffman, and M Gobbini. The distributed human neural system for face perception. *Trends Cogn Sci (Regul Ed)*, 4(6):223–233, Jun 2000.
- [29] F Jiang, L Dricot, V Blanz, R Goebel, and B Rossion. Neural correlates of shape and surface reflectance information in individual faces. *Neuroscience*, 163(4):1078–91, Nov 2009.
- [30] X Jiang, E Rosen, T Zeffiro, J VanMeter, V Blanz, and M Riesenhuber. Evaluation of a shape-based model of human face discrimination using fmri and behavioral techniques. *Neuron*, 50(1):159–172, 2006.
- [31] Wilsaan M. Joiner, Jung-Eun E. Lee, and Mark Shelhamer. Behavioral analysis of predictive saccade tracking as studied by countermanding. *Exp Brain Res*, 181(2):307–320, May 2007.
- [32] N Kanwisher, J McDermott, and M M Chun. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, 17(11):4302–11, Jun 1997.

- [33] R. Kawashima, J. Tanji, K. Okada, M. Sugiura, K. Sato, S. Kinomura, K. Inoue, A. Ogawa, and H. Fukuda. Oculomotor sequence learning: a positron emission tomography study. *Exp Brain Res*, 122(1):1–8, 1998.
- [34] Christina S Konen and Sabine Kastner. Two hierarchically organized neural systems for object information in human visual cortex. *Nat Neurosci*, 11(2):224–31, Feb 2008.
- [35] C Lee, W H Rohrer, and D L Sparks. Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature*, 332(6162):357–60, Mar 1988.
- [36] Gunter Loffler, Grigori Yourganov, Frances Wilkinson, and Hugh R Wilson. fmri evidence for the neural representation of faces. *Nat Neurosci*, 8(10):1386–1391, Oct 2005.
- [37] N K Logothetis, J Pauls, M Augath, T Trinath, and A Oeltermann. Neurophysiological investigation of the basis of the fmri signal. *Nature*, 412(6843):150–7, Jul 2001.
- [38] Alexander Maier, Melanie Wilke, Christopher Aura, Charles Zhu, Frank Q Ye, and David A Leopold. Divergence of fmri and neural signals in v1 during perceptual suppression in the awake monkey. *Nat Neurosci*, 11(10):1193–200, Oct 2008.
- [39] D P Munoz, M C Dorris, M Paré, and S Everling. On your mark, get set: brainstem circuitry underlying saccadic initiation. *Can J Physiol Pharmacol*, 78(11):934–44, Nov 2000.
- [40] D P Munoz and R H Wurtz. Saccade-related activity in monkey superior colliculus. i. characteristics of burst and buildup cells. *Journal of Neurophysiology*, 73(6):2313–33, Jun 1995.
- [41] L.A. Murotek, C.C.A.M. Gielen, and M. Flanders. Manual tracking in three dimensions. *Exp Brain Res*, 171:99–115, 2006.
- [42] W T Newsome and E B Paré. A selective impairment of motion perception following lesions of the middle temporal visual area (mt). *J Neurosci*.
- [43] S Ogawa and T M Lee. Magnetic resonance imaging of blood vessels at high fields: in vivo and in vitro measurements and image simulation. *Magn Reson Med*, 16(1):9–18, Oct 1990.

- [44] S Ogawa, T M Lee, A R Kay, and D W Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci USA*, 87(24):9868–72, Dec 1990.
- [45] F P Ottes, J A Van Gisbergen, and J J Eggermont. Visuomotor fields of the superior colliculus: a quantitative model. *Vision Res*, 26(6):857–73, Jan 1986.
- [46] M. Pavel. Predictive control of eye movement. *Rev Oculomot Res*, 4:71–114, 1990.
- [47] C Pierrot-Deseilligny, R M Müri, C J Ploner, B Gaymard, and S Rivaud-Péchoux. Cortical control of ocular saccades in humans: a model for motricity. *Prog Brain Res*, 142:3–17, Jan 2003.
- [48] M. K. Rand, O. Hikosaka, S. Miyachi, X. Lu, K. Nakamura, K. Kitaguchi, and Y. Shimo. Characteristics of sequential movements during early learning period in monkeys. *Exp Brain Res*, 131(3):293–304, 2000.
- [49] Morisaki Reiko and Inui Toshio. Neural network model of face recognition in the inferotemporal cortex. *IEICE technical report. Neurocomputing*, 598(95):151–158, Jan 1996.
- [50] D A Robinson and A F Fuchs. Eye movements evoked by stimulation of frontal eye fields. *Journal of Neurophysiology*, 32(5):637–48, Sep 1969.
- [51] D.A. Robinson, J.L. Gordon, and S.E. Gordon. A model of the smooth pursuit eye movement system. *Biol Cybern*, 55:43–57, 1986.
- [52] C D Salzman, C M Murasugi, K H Britten, and W T Newsome. Microstimulation in visual area mt: effects on direction discrimination performance. *J Neurosci*.
- [53] J D Schall. Neuronal activity related to visually guided saccadic eye movements in the supplementary motor area of rhesus monkeys. *Journal of Neurophysiology*, 66(2):530–58, Aug 1991.
- [54] J D Schall, D P Hanes, K G Thompson, and D J King. Saccade target selection in frontal eye field of macaque. i. visual and premovement activation. *J Neurosci*, 15(10):6905–18, Oct 1995.

- [55] J Schlag and M Schlag-Rey. Evidence for a supplementary eye field. *Journal of Neurophysiology*, 57(1):179–200, Jan 1987.
- [56] G. Schwartz. Estimating the dimension of a model. *The Annals of Statistics*, 6:461–464, 1978.
- [57] M. Shelhamer and W. M. Joiner. Saccades exhibit abrupt transition between reactive and predictive; predictive saccade sequences have long-term correlations. *J Neurophysiol*, 90(4):2763–2769, 2003.
- [58] T. Shibata, H. Tabata, S. Schaal, and M. Kawato. A model of smooth pursuit in primates based on learning the target dynamics. *Neural Netw*, 18(3):213–224, 2005.
- [59] L Sirovich and M Kirby. Low-dimensional procedure for the characterization of human faces. *J Opt Soc Am A*, 4(3):519–24, Mar 1987.
- [60] Mona Spiridon, Bruce Fischl, and Nancy Kanwisher. Location and spatial profile of category-specific regions in human extrastriate cortex. *Hum. Brain Mapp.*, 27(1):77–89, Jan 2006.
- [61] L. Stark, G. Vossius, and L.R. Young. Predictive control of eye tracking movements. *Institute of Radio Engineers Trans in Human Factors and Electronics*, 3:52–57, 1962.
- [62] Y Sugase, S Yamane, S Ueno, and K Kawano. Global and fine information coded by single neurons in the temporal visual cortex. *Nature*, 400(6747):869–73, Aug 1999.
- [63] J. Tanji and K. Shima. Role for supplementary motor area cells in planning several movements ahead. *Nature*, 371(6496):413–416, 1994.
- [64] Ivan Toni, Michael Krams, Robert Turner, and Richard E. Passingham. The time course of changes during motor sequence learning: A whole-brain fmri study. *NeuroImage*, 8(1):50–61, 1998.
- [65] Matthew Alan Turk, Alexander Pentland, Massachusetts Institute of Technology. Media Laboratory. Vision, and Modeling Group. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):16, Jan 1991.

- [66] LG Ungerleider and M Mishikin. Two cortical visual systems. in: Ingle dj, mansfield rjw, goodale ma (eds) advances in the analysis of visual behavior. *MIT Press, Cambridge Mass.*
- [67] T O Videen, J S Perlmutter, P Herscovitch, and M E Raichle. Brain blood volume, flow, and oxygen utilization measured with 15o radiotracers and positron emission tomography: revised metabolic computations. *J Cereb Blood Flow Metab*, 7(4):513–6, Aug 1987.
- [68] G. Westheimer. Eye movement response to a horizontally moving visual stimulus. *A.M.A. Archives of Ophthalmology*, 52:932–941, 1954.
- [69] S Yamane, S Kaji, and K Kawano. What facial features activate face neurons in the inferotemporal cortex of the monkey? *Experimental brain research Experimentelle Hirnforschung Expérimentation cérébrale*, 73(1):209–14, Jan 1988.
- [70] Hongchuan Zhang, Jiangang Liu, David E Huber, Cory A Rieth, Jie Tian, and Kang Lee. Detecting faces in pure noise images: a functional mri study on top-down perception. *Neuroreport*, 19(2):229–33, Jan 2008.