

エネルギー科学研究科
エネルギー社会・環境科学専攻修士論文
パフォーマンス-認知モデルを用いた
題目： 精神負荷作業中の認知・作業状態
推定手法

指導教員： 下田 宏 教授

氏名： 國政 秀太郎

提出年月日： 平成26年2月7日(金)

論文要旨

題目：パフォーマンス-認知モデルを用いた精神負荷作業中の認知・作業状態推定手法

下田研究室，國政 秀太郎

要旨：

近年，企業は電力を中心とするエネルギー消費量削減を推し進め，クールビズの実施や照明の間引き等の省エネルギー活動を積極的に行っている一方で，執務者の快適性への影響を軽視する傾向にあるため，執務環境が悪化し，執務者の知的生産性が低下する可能性が危惧されている．今後は省エネルギー化を促進させながら，知的生産性の低下を抑える執務環境の設計や改善が望まれるが，現在も未だ有効な知的生産性評価手法が確立されていない．これまでに提案されてきた評価手法では，認知タスクの単純な解答時間や正答率の評価が主流であるが，このタスクパフォーマンスの変動が執務環境によるものなのか，執務者の内的要因（認知負荷など）によるものかを評価し難いという問題が挙げられる．

そこで本研究では，パフォーマンスだけでなく認知負荷も定量的に評価し，それぞれを適切に解釈することで，執務者がどのような認知状態であるのかを評価する手法を提案した．具体的には，パフォーマンス及び認知負荷によって定義される認知・作業状態のモデルとして，パフォーマンス-認知モデルを提案し，ルールベースファジィ推論によるパフォーマンスおよび認知負荷の定量評価から，提案モデルを用いた執務者の認知・作業状態の時系列推定手法を提案した．また，本手法の妥当性を評価するために，難易度の異なる複数のタスク（2桁暗算及び4桁暗算）を用いた被験者実験を行い，認知・作業状態を推定する上でパフォーマンス評価および認知負荷評価の妥当性を確認した．

被験者実験の結果，提案手法において推定されたパフォーマンス及び認知負荷がタスク難易度上昇によってそれぞれ有意に低下及び上昇することが分かったため，提案手法による推定は妥当であると示唆された．さらに，同じ認知・作業状態であっても，前後の状態遷移から異なる解釈ができ，より詳細に執務者の作業状況を推測できる可能性が示唆され，より効果的な執務環境評価や改善案の検討に貢献できると考えられる．

目次

第 1 章 序論	1
第 2 章 研究の背景と目的	3
2.1 研究の背景	3
2.2 既往研究	4
2.2.1 知的生産性評価	4
2.2.2 Cognitive Load Theory 研究	5
2.3 研究の目的	7
2.4 研究の意義	7
第 3 章 パフォーマンス-認知モデルを用いた認知・作業状態推定手法	8
3.1 パフォーマンス-認知モデル	8
3.1.1 提案モデルと認知・作業状態の定義	8
3.1.2 モデルの適用例	10
3.2 認知・作業状態推定手法の概要	11
3.3 ルールベースファジィ推論	13
3.3.1 ファジィ推論と 2 値推論の違い	14
3.3.2 ファジィ集合とメンバシップ関数の定義	15
3.3.3 推論方法	17
3.3.4 推論プログラムのアルゴリズム	20
3.4 認知タスク	24
3.5 パフォーマンス推定	26
3.5.1 特徴量抽出	27
3.5.2 メンバシップ関数の設定	28
3.5.3 規則設定	34
3.6 認知負荷推定	36
3.6.1 特徴量抽出	36
3.6.2 メンバシップ関数の設定	40

3.6.3	規則設定	41
3.7	認知・作業状態推定	42
第 4 章	提案手法の妥当性評価実験	45
4.1	目的	45
4.2	実験方法	45
4.2.1	実験環境	45
4.2.2	測定手法	46
4.2.3	実験スケジュール	47
4.2.4	被験者	50
4.3	実験結果	52
4.3.1	NASA-TLX	53
4.3.2	解答時間, 分散性及びエラー率	53
4.3.3	心拍数及び瞳孔径	56
4.3.4	パフォーマンス、認知負荷、認知・作業状態の推定結果	58
4.4	考察	66
4.4.1	特徴量の検討	66
4.4.2	パフォーマンス推定及び認知負荷推定の検討	67
4.4.3	認知・作業状態推定の検討	68
4.4.4	今後の課題	71
第 5 章	結論	72
	謝辞	73
	参考文献	74

目次

2.1	認知負荷の時系列データ例 ^[8]	5
3.1	パフォーマンス-認知モデル	9
3.2	認知状態推定手法の概念図	12
3.3	ファジィ推論のブロック図	13
3.4	代表的なメンバシップ関数の形状	16
3.5	ガウシアン型メンバシップ関数の形状	17
3.6	規則 R^i での推論	18
3.7	複数の推論結果の統合	19
3.8	ファジィ推論のアルゴリズム	20
3.9	後件部のメンバシップ関数形状	22
3.10	タスクのインタフェースとタスク解答の流れ	25
3.11	入力特徴量に対するメンバシップ関数の形状	29
3.12	瞬目による瞳孔径アーチファクトの除去	37
3.13	タスクごとの特徴量抽出	38
3.14	心電図 (ECG) の例	39
3.15	認知・作業状態推定アルゴリズム	44
4.1	実験室のレイアウト	45
4.2	被験者と計測機材の位置関係	46
4.3	被験者 s06 の心電図データ	52
4.4	メンタルワークロードのグラフ及び検定結果	53
4.5	各タスクごとの平均解答時間	54
4.6	各タスクごとの平均分散性	54
4.7	各タスクごとの平均エラー率	54
4.8	各タスクごとの平均瞳孔径	56
4.9	各タスクごとの平均心拍数	56
4.10	各タスクごとの平均パフォーマンス	58

4.11 各タスクごとの平均認知負荷	58
4.12 解答時間, 分散性, エラー率及びパフォーマンス推定結果	59
4.13 瞳孔径, 心拍数及び認知負荷推定結果	59
4.14 High W.S. の平均適応度	61
4.15 Optimal W.S. の平均適応度	61
4.16 Bad W.S. の平均適応度	61
4.17 Low W.S. の平均適応度	61
4.18 被験者 s10 の認知・作業状態推定結果 (2桁暗算)	63
4.19 被験者 s10 の認知・作業状態推定結果 (4桁暗算)	64
4.20 被験者 s22 の認知・作業状態推定結果 (2桁暗算)	65
4.21 被験者 s22 の認知・作業状態推定結果 (4桁暗算)	66

表 目 次

3.1	2桁暗算タスクの認知ステップと最短および最長処理時間	31
3.2	4桁暗算タスクの認知ステップと最短および最長処理時間	32
3.3	解答時間の理論値	33
3.4	パフォーマンス推定での規則	35
3.5	認知負荷推定での規則	41
3.6	認知状態推定における規則	42
4.1	初日の実験スケジュール	47
4.2	事前アンケート項目	48
4.3	2日目の実験スケジュール	49
4.4	被験者属性	51
4.5	NASA-TLXによるメンタルワークロード	53
4.6	各タスクの平均特徴量(解答時間・分散性・エラー率)	54
4.7	各被験者の平均特徴量(解答時間, 分散性, エラー率)	55
4.8	各タスクの平均特徴量(瞳孔径・心拍数)	56
4.9	各被験者の平均特徴量(瞳孔径, 心拍数)	57
4.10	各タスクの平均パフォーマンス及び平均認知負荷	58
4.11	各被験者の平均パフォーマンス及び平均認知負荷	60
4.12	各タスクの認知・作業状態の平均適応度	61
4.13	各被験者の認知・作業状態の平均適応度	62

第 1 章 序論

現在の情報化社会では、かつて人間の行ってきた作業の多くをコンピュータが行うようになり、人間の遂行すべき作業は意思決定やアイデア創造などの、より高度な知的作業となってきている。特に、企業や研究機関において高次の知的作業の価値は年々高まってきており、経済的および社会的利益を得るためには知的作業の効率や正確さなどの知的生産性を向上させることが重要である。一方で、企業では照明の間引きなど、環境配慮や電力料金削減を目指した省エネルギー化を進めている。しかし、このような執務環境の見直しは作業執務者に対して不快な環境条件になる可能性があり、知的生産性の低下を招く恐れがある。このように、知的生産性を考慮せずにエネルギー消費削減の観点のみから執務環境を見直すことには問題があり、今後はエネルギー消費量だけでなく、知的生産性への影響も評価しながら執務環境を改善すべきである。しかし、現在も有効な知的生産性評価手法は確立されていない。したがって、執務環境改善のためには、知的生産性評価手法の確立が急務であり、これまで多くの研究が行われてきた^[1-6]。これらの研究は、執務者の知的生産性を評価する際、実験用に設計されたタスクを使用し、その作業記録から作業速度や正答率などのタスクパフォーマンスを評価している。

しかし、知的生産性の変動は執務環境の照明や空気質などの外的要因だけでなく、執務者にかかる認知負荷などの内的要因にも影響されるため、作業速度や正答率の単純な比較のみによる知的生産性の評価手法では、ある執務環境下で知的生産性の変動した時、その変動が外的要因によるものなのか内的要因によるものなのか評価ができない。外的要因と知的生産性の関係についての研究は、Wargock ら^[4]による空気質と知的生産性の関係についての研究などが挙げられるが、客観的な内的要因の評価は行われていないため、外的要因と知的生産性の関係については未だ議論が残っている。したがって、今後は知的生産性評価だけでなく、内的要因の客観評価も必要である。一方で、人間の認知活動に強く影響を与えると考えられている認知負荷についての研究も数多く行われており^[7-12]、アンケートなどの主観指標だけでなく、瞳孔径や心拍数などの生理指標から客観的に評価できることから、認知負荷も考慮した新しい知的生産性評価手法の確立が可能になると期待されている。

そこで本研究では、タスクパフォーマンスと認知負荷によって決定される執務者の認

知・作業状態を仮定したパフォーマンス-認知モデルを提案し、作業執務者のパフォーマンスと認知負荷の推定、さらに認知・作業状態を推定する手法を開発した。パフォーマンス推定は認知タスクの作業記録から行い、認知負荷の推定では瞳孔径および心拍数を用いた。各推定にはルールベースのファジィ推論を利用し、過去の知見も利用したルール設定を行った。この提案手法が確立することで、執務者の認知・作業状態の評価が可能となり、各状態の遷移状況や全作業時間に占める割合から、執務者の作業状況が詳細に評価できる。これにより、執務者にとって快適な執務環境設計や改善がより正確かつ簡便になると期待される。

本論文は、第1章の序論を含め5章で構成されている。第2章では、研究の背景として、環境配慮や省エネ対策による執務環境の見直しが知的生産性に与える影響や知的生産性評価の重要性を述べた後、既往研究として知的生産性研究及び Cognitive Load Theory 研究について説明する。第3章では、認知・作業状態推定手法を提案し、その詳細について述べる。第4章では、提案手法の妥当性評価実験について述べる。第5章では、本研究の結果をまとめ、今後の課題を述べる。

第 2 章 研究の背景と目的

本章では、近年の環境配慮および省エネ対策によるオフィス環境見直しと知的生産性の関係について述べ、関連研究としてこれまでの知的生産性評価と Cognitive Load Theory 研究について説明し、問題点を挙げた上で本研究の目的と意義を述べる。

2.1 研究の背景

3.11 東日本大震災以来、東京電力福島第一原子力発電所の事故から始まった慢性的な電力不足により、企業は電力を中心とするエネルギー消費量削減を推し進め、クービズの実施や照明の間引き等の省エネルギー活動を積極的に行っている。しかし、エネルギー消費量削減を重視する一方で、執務者の快適性への影響を軽視する傾向にあるため、執務環境が悪化し、執務者の知的生産性が低下する可能性が危惧されている。情報化が進んだ現代において、オフィスにおける多くの作業が書類作成、情報管理および処理という知的作業であるため、執務者の知的生産性の低下は業務へ支障をきたすばかりか、労働時間の増加ひいては人件費の増加につながる恐れがある。一般に企業の支出の中で人件費は大きな割合を占めており、執務環境に対する光熱費や設備投資費に比べて高い。したがって、コスト削減や省エネルギーのためにエネルギー消費量削減を実施した結果、執務環境の悪化による残業の増加等が人件費やエネルギー消費量の増加を招き、かえって支出が拡大するばかりか、省エネルギー活動を阻害する危険性がある。以上より、エネルギー消費量削減にのみ着目するのではなく、執務者の知的生産性も検討し、執務環境の設計や改善を行う必要がある。執務環境の変化による知的生産性の変動を評価できれば、エネルギー消費量を削減しながら知的生産性の低下を抑える執務環境の設計が可能になるだけでなく、知的生産性を向上させる執務環境の設計も可能となる。執務者の知的生産性が向上した場合、その向上率が数%であったとしてもその金額換算値は大きく、環境改善による費用対効果は高いという報告もある^[13]ため、知的生産性評価が可能になる意義は大きい。したがって、その評価手法の確立は急務であるが、現在も未だ有効な評価手法が確立されていない。

次節では以上のような研究背景に関連して、知的生産性の評価手法や認知負荷などの内的要因の評価を目指した既往研究を挙げる。

2.2 既往研究

2.2.1 知的生産性評価

研究者によって知的生産性の定義は様々であり^[14]，村上は知的作業の効率が知的生産性であると示唆している^[15]．また同様に，知的作業の定義も様々であるが，村上らの知的活動の階層モデル^[16]によると，知的作業は(1)知覚認知や判断・操作などの「情報処理」と，(2)知識情報の記憶や計算などの「知識処理」，(3)価値創造やイノベーション，閃きなどの「知識創造」の3階層に分けられるとされ，村上らはオフィスやオペレーション業務などの様々な知的作業は「情報処理」や「知識処理」を扱うものが多いと示唆している．

現在の高度情報化が進んだ社会において，人間が行う作業の多くが知的作業であり，知的生産性の重要度は高まっている．したがって，知的生産性向上を目指した環境設計などの取り組みが行われるべきであるが，その評価手法は未だ確立されていない．特に，その定量評価は困難であり，近年の知的生産性研究はこの問題に焦点が当てられている．Wargockiら^[4]は，汚染物質や換気量などの空気質とテキストタイピング等の作業成績との関係を明らかにしている．また，知的生産性評価のためのタスク開発も様々な研究で行われており，DavidらはPAB^[5]と呼ばれる脳の高次機能を評価するタスクを開発している．同様に下田らはタスクセットCPTOP^[2]，岩川らはCPTOP2^[3]を開発しており，知的生産性の定量評価ツールの確立を目指している．これらの開発されたタスクセットを用いることで，知的生産性向上を目指した執務環境設計が可能になると期待される．しかし，これらの研究ではタスクの解答速度や正答率などの作業成績を評価しているが，このような執務者のタスクパフォーマンスのみからの知的生産性評価は新しい執務環境の設計や改善を行うにあたって不十分であると考えられる．その理由の一つとして，執務者の内的要因が考慮されていないことが挙げられる．執務者の心理状態および執務者にかかる認知負荷のような内的要因が評価されていないければ，ある執務環境下で作業成績が変動した場合，それが執務者の内的要因によるものか，もしくは照明や空気質などの外的要因によるものなのかという評価ができない．したがって，知的生産性評価は執務者に関する様々な情報から行われるべきであり，作業成績などのタスクパフォーマンスだけでなく，心理状態や認知負荷などの内的要因の検討が必要であるが，現在の知的生産性研究では，後者も考慮した研究が非常に少ない．

2.2.2 Cognitive Load Theory 研究

人間の認知状態などの内的要因に着目した研究は多く存在し、特に知的能力に関係のある作業記憶や認知負荷に深い繋がりを持つ Cognitive Load Theory (CLT) 研究^[7-12]が挙げられる。これらの研究は人間の認知負荷を測定することで人間の内的状態の特定を目指し、主に教育の分野で、学習者を対象に実施されてきた。Fred^[8]らによると、認知負荷は大きく分けて intrinsic load, germane load, extraneous load の3つの種類に分けられ、知的活動を行う際は図2.1のようにそれぞれが積み重なった負荷を受けていると示唆している。

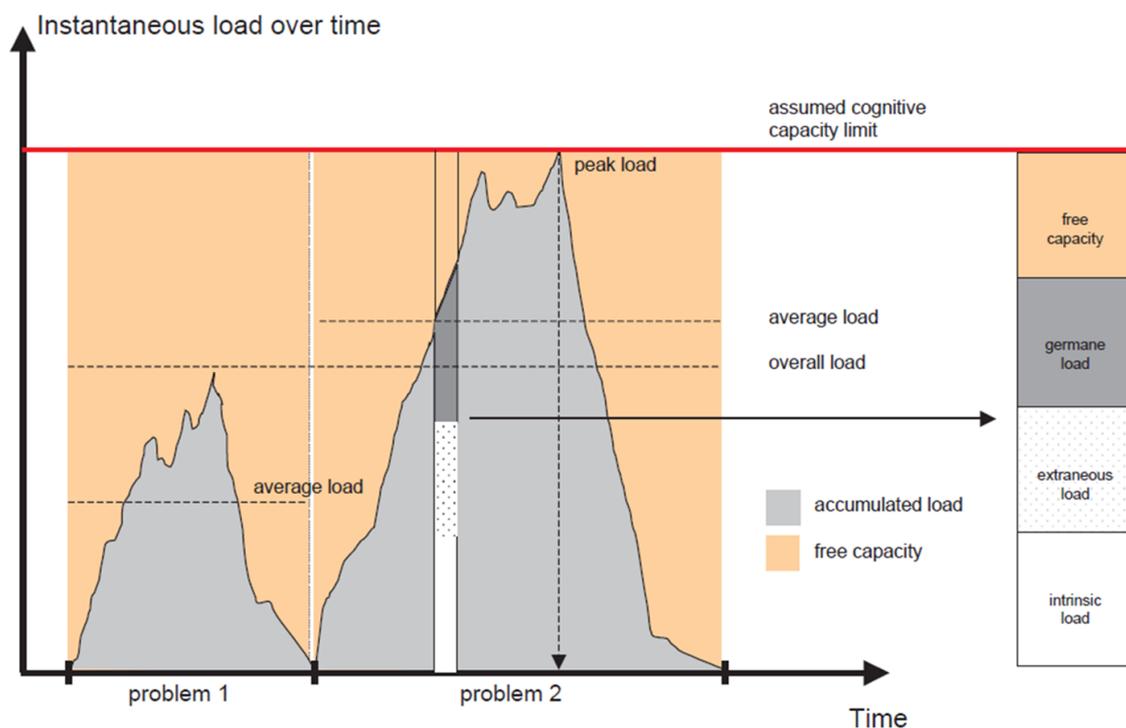


図 2.1: 認知負荷の時系列データ例^[8]

人間の認知資源つまり認知負荷容量は有限であり、図中の赤い横線がその限界容量である。ある知的活動の認知負荷が限界容量を超えなければ、高いパフォーマンスが期待できるが、認知負荷が限界容量を超える場合、十分にパフォーマンスが発揮されないと考えられている。ここで intrinsic load とは学習者に潜在的にある要因 (先行知識など) に依存する負荷であり、germane load とは学習に直接関係するメンタルな要因 (モチベーションなど) に依存する負荷、extraneous load とは学習に直接関係しない阻害要因 (テキストの見づらさなど) に依存する負荷である。intrinsic load は初学者や困難な課題に対しては高くなる傾向があり、逆に熟達者や容易な課題に対しては低くな

る傾向があるが、学習者自身の知識などに依存して変動するため、一般に教育デザインによるコントロールが困難な負荷である。一方、他の2種類の負荷に関してはコントロールがある程度可能とされており、最適な教育デザイン下では *germane load* は高く、*extraneous load* は低いとされている。特に前者はモチベーション等に関わる負荷であるので、学習においてこの負荷が高い時、効率的に学習が進むと言われている。また、Paas ら^[8,9]によると、作業成績などのパフォーマンスと認知負荷の2指標を同時に評価することが、人間の知的活動の理解にとって重要であると示唆されている。

認知負荷の評価については、生理指標を用いる研究も数多くあり、その有効性も示されている。Hess ら^[17]は瞳孔径と単純な認知タスク解答による知的活動との相関を示唆し、Stanners ら^[18]は瞳孔径は覚醒と認知活動を反映する指標として有効であることを示唆している。また、山本ら^[19]によって数列記憶タスクで瞳孔面積(瞳孔径)と瞬目時間に有意な変化が見られることが示されている。Halverson ら^[20]は瞳孔径だけでなく、瞬目頻度やサッケード眼球運動頻度、閉瞼率(*eye closure*)や閉瞼時間(*percentage of eye closure, PERCLOS*)を入力として、被験者の行ったタスクの異なる難易度による認知負荷の高低を判別することが可能であるか実験を行っている。結果として、瞳孔径やPERCLOSが特に判別能力の高い指標であると示唆されている。また、Iqbal ら^[21]はルート計画タスクやテキスト編集タスクなどの高次認知タスクを細かなステップ(サブタスク)に分け、それぞれの区間内での瞳孔径から認知負荷を評価している。

認知負荷を反映するとされている指標は瞳孔径だけでなく、心拍などの循環器系も挙げられている。石橋ら^[22]は心拍数が精神負担の指標として有効であることを示唆しており、Turner ら^[23]は暗算タスクやビデオゲーム、Douglas ら^[24]はRaven's matricesと呼ばれる知能検査、Mulder ら^[25]は文章校正タスクを用いてタスク時に心拍数が上昇することを確認している。特にTurnerらは認知タスクによる心拍上昇は、酸素消費量の上昇によるものであると示唆している。また、身体運動量は同じであっても、記憶想起タスクにおいてより高い認知負荷のタスクの方がより高い心拍数が計測されることを示唆した研究^[26]も存在している。これらの研究を受けて、Kennedy ら^[27]は認知負荷計測に心拍数を用いて、グルコース入り飲料とプラセボ飲料服用時でタスクパフォーマンスの相違を評価しており、前者でパフォーマンスが上昇することを示唆している。

以上のように、CLT研究は盛んに行われ、特に生理指標を用いた評価は有効であり、教育デザインや作業環境の設計に貢献してきた歴史がある。したがって、同様に知的活動を行う執務者の知的生産性評価においても、認知負荷を考慮することは有効であ

ると言える。しかし、CLT 研究や知的生産性研究においてパフォーマンスもしくは認知負荷の一方のみを評価する研究が数多く行われているものの、その両方を考慮した研究は少ない。

2.3 研究の目的

以上の背景より、知的生産性評価ではパフォーマンスだけでなく認知負荷の評価も必要であるが、何らの手法でそれぞれの評価値を出力しただけでは、執務環境の設計や改善には繋がらず、2 指標を適切に解釈し、執務者がどのような認知状態であるか評価する必要がある。認知状態を評価することが可能となれば、執務者の作業状況 (集中、休息など) が評価できるため、例えば「改善後の執務環境は作業に集中しやすい環境もしくは休息が適度に行われる環境である」などという、従来の手法では困難であった新しい執務環境の設計や改善が行えると考えられる。そのためにはまずパフォーマンスと認知負荷の 2 指標を用いた新しい認知モデルを提案する必要がある。

そこで本研究では、パフォーマンスおよび認知負荷の 2 軸で構成されるパフォーマンス-認知モデルを提案し、執務者の認知・作業状態の推定を目指す。さらにコンピュータ推論によるパフォーマンスおよび認知負荷の定量評価から、提案モデルを用いた執務者の認知・作業状態の推定手法を提案する。また、本手法の妥当性を評価するために、難易度の異なる複数のタスクを用いた被験者実験を行い、認知・作業状態を推定する上でパフォーマンス評価および認知負荷評価の妥当性を確認することを目的とする。

2.4 研究の意義

本研究の提案手法が確立すれば、執務者の複数の認知・作業状態を評価することが可能となり、各状態の遷移状況や全作業時間に占める割合から、より詳細な作業状況が推定できる。これにより、執務者にとって快適な執務環境設計や改善がより正確かつ簡便に行えるようになると期待される。また、本手法は執務者の認知タスク遂行と並行して認知・作業状態のリアルタイム推定が可能となるよう設計されているため、認知作業状態の時系列解析が可能となり、作業従事中の執務者の状態の経時変化や、静的とはいえ難い環境要因 (照度変動する照明環境、非周期的に発生するノイズ環境など) における知的生産性への影響評価が可能になると期待される。

第 3 章 パフォーマンス-認知モデルを用いた認知・作業状態推定手法

本章では、まずパフォーマンス-認知モデルを提案し、そのモデルを用いた認知・作業状態推定手法を述べる。その際、本手法に用いたルールベースファジィ推論について説明した後、この推論によるパフォーマンス推定および認知負荷推定について述べる。最後に、認知・作業状態の推定手法について説明する。

3.1 パフォーマンス-認知モデル

3.1.1 提案モデルと認知・作業状態の定義

従来の知的生産性研究では、タスクの作業記録(解答速度など)を用いた評価手法が主流であった。しかし、執務者の認知状態や認知負荷などの内的要因の評価は従来手法では考慮されておらず、人間が知的作業を行っている以上、作業への集中度やモチベーションの影響は無視できない。これらはパフォーマンスや認知負荷に影響するため、パフォーマンスだけでなく認知負荷も考慮する必要がある。この2指標を用いることで執務者の認知・作業状態の評価が新たに可能となる。例えばパフォーマンスが低い場合、認知負荷が低い場合と高い場合で執務者の認知・作業状態の評価は異なる。前者では休息していると考えられ、後者では集中の乱れなどが原因で作業遂行が困難になっていると考えられる。これら2状態は明らかに異なり、パフォーマンス評価のみでこの状態の切り分けは困難である。このような認知・作業状態の切り分けは、以上の2指標をそれぞれ高低の2パターンに分けて評価することで可能となるため、先述の例に挙げたように明らかに異なる状態を分類でき、ある状態への遷移を促したりもしくは抑えたりするような執務環境設計や改善が可能となる。本研究では執務者の認知・作業状態を4つに分類し、図3.1に示すパフォーマンス-認知モデルとして提案する。

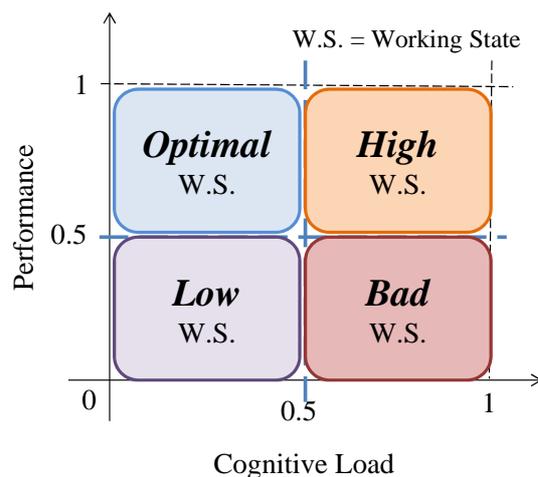


図 3.1: パフォーマンス-認知モデル

High Working State (High W.S.)

認知負荷が高く、かつ高いパフォーマンスが発揮されている状態である。この状態下にある執務者は作業に集中していると考えられ、この状態が多く観測される場合は良い作業成績が期待されるが、認知負荷が高いため、疲労が蓄積される。

Optimal Working State (Optimal W.S.)

執務者が低い認知負荷で高いパフォーマンスを発揮している状態であり、作業効率の高い状態と言える。作業への習熟などによって作業が快調に進んでいるときに多く観測されると考えられる。この状態は High W.S. とは違い、認知負荷が低いため、High W.S. より疲労の蓄積が少ないと思われる。

Low Working State (Low W.S.)

認知負荷が低く、パフォーマンスも同様に低い状態である。この時、執務者は作業中断や休息するなどして認知資源を多くは割いていないと考えられる。この状態が観測されている時、作業成績は悪いが認知負荷が低いため疲労やストレスが解消され、後に High W.S. や Optimal W.S. に移行する可能性がある。

Bad Working State (Bad W.S.)

パフォーマンスが低いのに関わらず認知負荷が高い状態である。この状態下の執務者は作業以外の活動に認知資源を割くなどして作業が滞っているため、認知負荷が高い状態でありながらパフォーマンスが低下していると考えられる。さらに、Low W.S. と違い、疲労やストレスの解消が期待できない。

これら4状態が時系列で推定可能となった場合、それらの発生頻度や全作業時間に対する占有率も導出でき、より詳細な知的生産性評価が可能となる。

3.1.2 モデルの適用例

本モデルを用いた認知・作業状態の評価手法によって、より詳細な作業状況や知的生産性の評価が可能になると期待される。本モデルの適用例として様々な執務環境設計や改善が挙げられるが、ここでは執務環境を3つ想定し、本モデル適用の有用性を述べる。具体的には、(1) オフィス環境、(2) 学校授業などの教育環境、(3) 原子力発電プラントの監視室や航空管制室などのビジランス作業環境を想定する。これら3つの環境を想定した理由として、(1) および(2) はそれぞれ知的生産性研究と CLT 研究でよく想定される環境であること、(3) は、その執務環境の改善が重要であることが挙げられる。(3) のようなビジランス作業環境下では常に監視業務を行うため認知負荷が高く、判断ミスや遅延が経済的損失だけでなく、人的被害など重大なリスクを引き起こす可能性があるため、その執務環境の改善は常に必要とされている。

(1) オフィス環境

オフィス環境下では、終日業務を行うことが多いため、休憩のタイミングの検討や疲労の蓄積を抑える執務環境設計は重要である。例えば、長時間 High W.S. が継続的に観測された場合、作業は滞りなく進んでいるが疲労が蓄積している可能性がある。その後、もし Bad W.S. が観測されたならば、この状態への遷移は疲労の蓄積が進んだ結果であると考えられる。したがって、以上のような状態遷移がよく観測される場合は、休憩のタイミングを早める、休憩の回数を増やすなどの改善案が考えられる。また、業務時間内で Low W.S. が観測された場合、それが短時間のみ継続し、のちに Optimal W.S. または High W.S. に遷移しているのであれば、執務者が適切な小休憩を取ったとし、執務環境として問題はないと評価できる。逆に Low W.S. が長時間継続していた場合は、眠気やモチベーションの低下、過度な休憩などが生じていると解釈し、眠気などの防止策を講じる必要があると評価できる。

(2) 教育環境

学校の授業などの教育環境下では、短い時間で授業が区切られているため、生徒の学習効率を上げることが必要である。短時間で最大限の学習効果を期待する場合、Optimal W.S. や High W.S. が多く観測されることが望ましいのは勿論である

が、オフィス環境とは違い、Low W.S. については短時間であっても観測されるべきではないと評価される可能性も考えられる。Low W.S. や Bad W.S. が観測される場合は授業に集中しやすい環境への改善が求められる。

(3) ビジランス作業環境

原子力発電プラントや航空管制の監視作業を行うビジランス作業環境では、オペレーションのミスが大きな事故に繋がる危険性があるため、高いパフォーマンスが要求される。緊急時や現場の従業員への通信時などでは高い認知負荷がかかるが、それらの業務以外では単調な監視作業を行うため、退屈さによる覚醒度低下が問題になっており、高いパフォーマンスが要求される場面で十分なパフォーマンスが発揮されない可能性がある。したがって、このようなビジランス作業環境では、Bad W.S. だけでなく Low W.S. も危険性の高い状態であると評価される。また、High W.S. が長時間観測される場合、疲労が蓄積すると思われるため、後に Bad W.S. もしくは Low W.S. へ遷移する可能性がある。その場合は、オペレータに休憩を取らせるなどの対策が必要である。

以上のように、それぞれの環境下での認知・作業状態に対する評価や解釈は異なるが、パフォーマンスおよび認知負荷の2指標を用いた本モデルを用いて、各4状態を時系列で推定できることで、先述のように詳細な執務環境評価が可能になり、執務環境改善に向けた効果的な対策が実施可能となる。

3.2 認知・作業状態推定手法の概要

本研究で提案する認知・作業状態推定アルゴリズムは大別して3つの推定器モジュールで構築されている。状態推定の概念図は図3.2に示す通りであり、瞳孔径、心拍数、認知タスクの解答時間及び正誤(0 or 1のダミー変数)を入力とし、(1) 認知タスクの作業記録(解答時間及びエラー率)からのパフォーマンスの推定および定量化、(2) 瞳孔径及び心拍数からの認知負荷の推定および定量化、(3) パフォーマンスおよび認知負荷からの認知状態推定、以上3つによって本手法は構成されている。

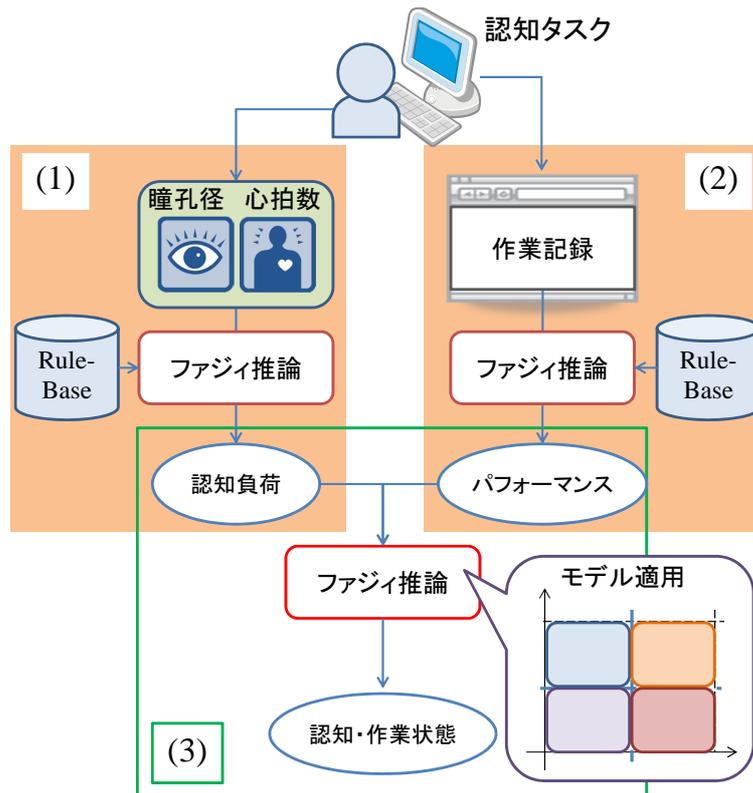


図 3.2: 認知状態推定手法の概念図

各推論にはルールベースファジィ推論を採用し，既往研究の知見に基づいてルールを設定し，認知状態を推定する．この推論手法を用いた理由として，この手法はあいまいさの概念を取り入れているため，パフォーマンス及び認知負荷のように明確にその高低を決定できる閾値が設定できない場合でも，ファジィ集合を導出し，過去の定性的な知見や情報をルールとして設定することで認知・作業状態が推定できることが挙げられる．

以下では，まず，3.3節で3つの推定器の基本となるルールベースファジィ推論について説明し，3.5節，3.6節ではそれぞれパフォーマンス推定および認知負荷推定でファジィ推論に用いる入力引数やメンバシップ関数の各種パラメータ，ルール(規則)設定について説明する．最後に3.7節では，定量化されたパフォーマンスおよび認知負荷から認知・作業状態を推定する手法を述べる．

3.3 ルールベースファジィ推論

執務者のパフォーマンス、認知負荷を推定するために、本研究ではルールベースのファジィ推論を採用した。ファジィ推論とはコンピュータ推論の1つであり、推論アルゴリズムにあいまいさの概念、つまりファジィ理論を適用させた推論である [28]。自然界、人間社会などに存在する様々なデータから行われる推論は、その結果がクリस्पに決まることは少なく、多くの場合がファジィ的である。ファジィ推論では、このようなあいまいな結果をファジィ集合およびファジィ論理によって導出する。本研究の推定対象であるパフォーマンスや認知負荷について、明確にその高低を決定する閾値は設定できず、既往研究にもその値を示唆できた例は少ない。したがって、本研究では先述のようなあいまいな結果を許容するファジィ推論を採用した。さらにこの推論法では、あらかじめルールを設定することが可能であり、これにより過去の定性的な知見や情報を利用できる。これはファジィ推論の利点の1つであり、本研究の提案手法でファジィ推論を採用した最大の理由でもある。

ファジィ推論は図3.3のように、(1)まず複数の入力データにファジィ化 (Fuzzification) と呼ばれる処理をすることで、ファジィ変数を複数個生成する。(2)次に、事前に設定されたルールベースに基づき、複数のファジィ変数から推論結果を示すファジィ集合を推論器 (Inference Engine) によって出力し、それらが数値化された結果を算出する非ファジィ化 (Defuzzification) と呼ばれる処理を行う。

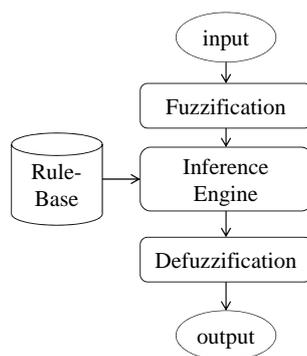


図 3.3: ファジィ推論のブロック図

次項以降では、まずファジィ推論と一般の2値推論の違いについて3.3.1項で述べ、上記の(1)のファジィ化や(2)の非ファジィ化を行うにあたってのファジィ集合やメンバーシップ関数の定義を3.3.2項で述べ、ファジィ推論の推論方法についての詳細を3.3.3

項で述べ、ファジィ推論の理論について説明する。最後に3.3.4項ではこの理論をプログラム実装した際のアゴリズムについて述べる。

3.3.1 ファジィ推論と2値推論の違い

コンピュータ推論は数多く存在するが、そのほとんどの基本は3段論法のような演繹推論であり、以下の例のような規則に従って行う。

前提 1 If x is A and y is B then z is C

前提 2 x is A and y is B

結論 z is C

一般の2値論理における推論では、 A 、 B および C は通常集合を扱うため、 x is A 、 y is B および z is C は $x \in A$ 、 $y \in B$ および $z \in C$ であり、特にIF文を前件部、THEN文を後件部と呼ぶ。また、それぞれの特性関数 $\mu_A(x)$ 、 $\mu_B(y)$ および $\phi_C(z)$ は以下のようなになる。

$$\mu_A(x) = \begin{cases} 1 & (\text{if } x \in A), \\ 0 & (\text{if } x \notin A), \end{cases} \quad (3.1)$$

$$\mu_B(y) = \begin{cases} 1 & (\text{if } y \in B), \\ 0 & (\text{if } y \notin B), \end{cases} \quad (3.2)$$

$$\phi_C = \min(\mu_A(x), \mu_B(y)) \quad (3.3)$$

ここで式3.3で用いた \min 関数は最小値を返す関数であるが、これはand演算つまり共通集合の関数である。2値論理の場合、 $\phi_C = 0$ であれば $z \notin C$ であり、 $\phi_C = 1$ であれば $z \in C$ となる。

ファジィ推論では、集合 A 、 B および C をファジィ集合として扱う。このとき、式3.1や式3.2の関数は0もしくは1の2値をとるのではなく、 $\mu_A(x) = f(x)$ のように $[0,1]$ をとる関数となる。この関数はメンバシップ関数と呼ばれ、集合 A における x の帰属度を表す。この値が1に近ければ、集合に属する度合が大きく、反対に0に近ければ度

合が小さいことを示している。特に、規則の前件部つまり集合 A や集合 B に対するメンバシップ関数を、前件部メンバシップ関数と呼び、後件部つまり集合 C に対するメンバシップ関数を後件部メンバシップ関数と呼ぶ。それぞれの関数形状やパラメータは推論対象 (パフォーマンス, 認知負荷など) によって異なるチューニングが施されるため、一意には決まらない。また、前件部メンバシップ関数によって得られた値は一般にファジィ変数と呼ばれ、この変数を導出することをファジィ化と呼ぶ。一方、 ϕ_C は前件部がどの程度規則に適応しているかの度合いを表し、さらに z の C に対する帰属度として解釈できる。最後にファジィ推論ではこの帰属度から後件部メンバシップ関数を用いて推論結果を出力し、非ファジィ化と呼ばれる推論結果を定量化する処理を行う。

次項ではファジィ集合およびメンバシップ関数の定義について述べる。

3.3.2 ファジィ集合とメンバシップ関数の定義

ファジィ推論を行うには、(1) 前件部および後件部に対するファジィ集合の定義、(2) メンバシップ関数の定義が必要である。ファジィ集合はあいまいな概念を表すものとして定義され、たとえば「赤い」や「重い」などである。定義するファジィ集合が多ければ、あいまいさを残しながらも詳細な推論が可能となるが、用意しなければならない規則数が各入力データに対する前件部ファジィ集合数の積となるため、規則数が膨大になる恐れがある。たとえば、入力データ 3 種類に対してそれぞれ 3 集合を定義した場合は $3 \times 3 \times 3 = 27$ 個の規則を用意すれば良いが、それぞれ 5 集合を定義した場合は 125 個用意しなければならない。一方でファジィ集合が少すぎる場合は、推論結果が漠然としたものとなるため推論を行う意義が失われる。

本研究では、定義する前件部のファジィ集合数を 3 とし、それぞれを「High」「Middle」「Low」とした。理由として (1) 計算負荷の低減と (2) 推論の細かさ維持の 2 点が挙げられる。(1) については、パフォーマンス推定および認知負荷推定に用いる入力データは 3.5 節や 3.6 節で後述するが、2 ~ 3 個であり、3 つのファジィ集合であれば用意しなければならない規則数は高々 27 個となるため計算量は低く、コンピュータの計算負荷の低減に繋がり、2.4 節に述べたようにリアルタイムの推論が可能である。(2) については、入力データを高低の 2 集合で評価するだけでなく中位の「Middle」の集合でも評価することで、ある程度推論結果の粗雑さを抑えることが可能であると考えた。同様の理由で、後件部のファジィ集合数も「High」「Middle」「Low」の 3 集合とした。

ファジィ集合を定義した後、それぞれの集合に対してメンバシップ関数を定義する

必要があるが、よく用いられている関数は3つあり、その関数形状は図 3.4 に示す通りである。

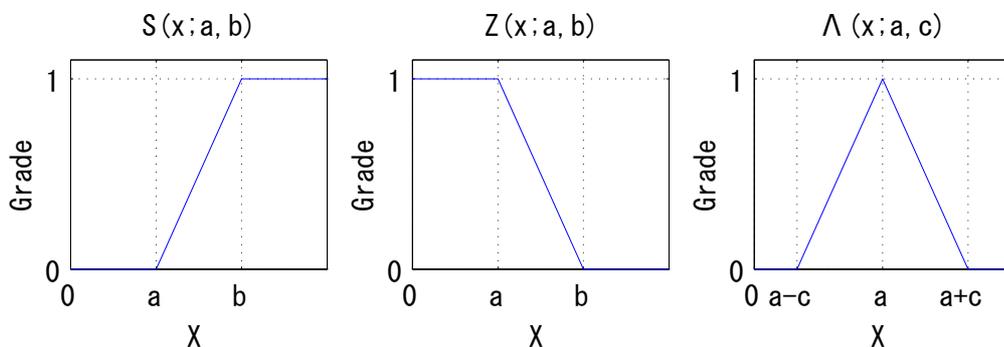


図 3.4: 代表的なメンバシップ関数の形状

また、各関数の数式は以下の通りである。(1) S 型関数

$$S(x; a, b) = \begin{cases} 0 & (\text{if } x < a), \\ \frac{x-a}{b-a} & (\text{if } a \leq x \leq b), \\ 1 & (\text{if } x > b), \end{cases} \quad (3.4)$$

(2) Z 型関数

$$Z(x; a, b) = \begin{cases} 1 & (\text{if } x < a), \\ \frac{b-x}{b-a} & (\text{if } a \leq x \leq b), \\ 0 & (\text{if } x > b), \end{cases} \quad (3.5)$$

(3) Λ 型関数

$$\Lambda(x; a, b) = \begin{cases} 1 - \frac{1}{c}|x-a| & (\text{if } |x-a| \leq c), \\ 0 & (\text{if } |x-a| > c), \end{cases} \quad (3.6)$$

ここで、 x は入力変数、 a および b はメンバシップ関数の閾値であり、 c はメンバシップ関数の幅である。入力変数の特性や値域を考慮した上であらかじめ設定しておく必要のある値である。

S 型や Z 型関数は入力 x が、ある閾値の外側では縦軸のグレードが 0 や 1 に固定される関数である。これらの関数は、入力値が増加するに従ってグレードが増加あるいは減少するようなファジィ集合 (Low, High) に対して設定するのに適している。対し

て Λ 関数は $x = a$ について対称な幅 c を持つメンバシップ関数であり、グレードが 1 をとる点が $x = a$ のみである。この関数は、入力値がある点から離れるに従ってグレードがその距離にのみ依存して変化するようなファジィ集合 (Middle) に対して設定するのに適している。一般にこれらの関数はファジィ推論でよく用いられる関数であるが、グレードが 0 もしくは 1 になる閾値が知られておらず、明確に決められない場合は、上記の関数形状をガウシアン型にして、あえてコンパクト・サポートのない関数を用いることがある。例えば Λ 型関数の場合、関数形状は図 3.5 のようになる。

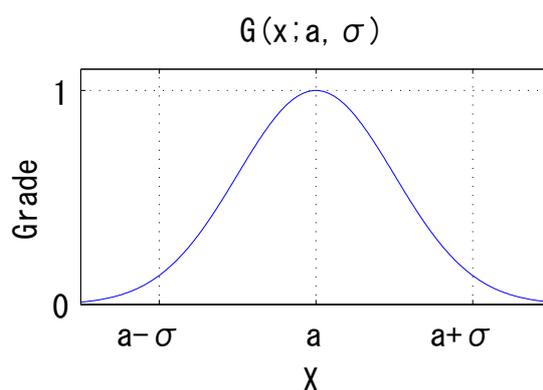


図 3.5: ガウシアン型メンバシップ関数の形状

また、数式は以下の通りである。

$$G(x; a, \sigma) = \exp\left\{-\frac{2(x-a)^2}{\sigma^2}\right\} \quad (3.7)$$

x は入力変数、 a と σ はそれぞれ関数の頂点の x 座標および関数の幅である。

パフォーマンス推定や認知負荷推定におけるファジィ化の閾値について、既往研究で明確に示唆された例がないため、本研究ではメンバシップ関数の形状は基本的にガウシアン型を用いた。閾値などのパラメータについては、パフォーマンス推定、認知負荷推定、認知・作業状態推定でそれぞれ異なるため後述する。

3.3.3 推論方法

実際に行われるファジィ推論では、3.3.1 項で述べたように単純な 1 つの推論ルール (以下では規則と呼ぶ) のみを用いることはなく、ファジィ化によって生成したファジィ変数を複数の規則へ適用し、それぞれの規則から推論を行い、各推論結果をまとめて数値化 (非ファジィ化) する。

ここで、 m 個の入力データ $x_k (k = 1, \dots, m)$ 及び 1 個の出力データ z に対して、 n 個のファジィ規則 $R^i (i = 1, \dots, n)$ が、次のように OR 結合されている場合について考える。

R^1 : if x_1 is A_1^1 and x_2 is A_2^1 and \dots and x_m is A_m^1 then z is C^1

or

\vdots

R^i : if x_1 is A_1^i and x_2 is A_2^i and \dots and x_m is A_m^i then z is C^i

or

\vdots

R^n : if x_1 is A_1^n and x_2 is A_2^n and \dots and x_m is A_m^n then z is C^n

ファジィ集合 A_k^i にはそれぞれ「High」「Middle」「Low」のいずれかが該当する。

ここで i 番目の規則について、前件部のメンバシップ関数 $\mu A_k^i(x_k)$ および後件部の特性関数 $\phi C^i(z)$ の関係式は式 3.3 と同様に、

$$\phi C^i = \min(\mu A_1^i(x_1), \mu A_2^i(x_2), \dots, \mu A_m^i(x_m)) \quad (3.8)$$

である。いま、入力 x_1^0, \dots, x_m^0 が与えられたとき、規則 R^i の推論結果を表すファジィ集合を C_*^i とし、 C^i が持つメンバシップ関数を $\mu C^i(z)$ とすると、上式から図 3.6 のように推論結果を出力できる。簡単のために図中では $m = 2$ とした。

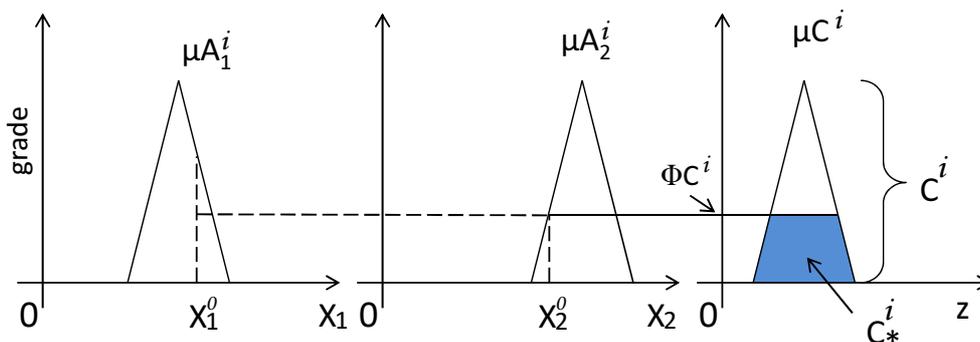


図 3.6: 規則 R^i での推論

このとき規則 R^i は $i = 1 \sim n$ で OR 結合されているので、それぞれを統合した推論結果 C^0 は

$$C^0 = C_*^1 \cup C_*^2 \cup \dots \cup C_*^n = \bigcup_{i=1}^n C_*^i \quad (3.9)$$

となり， n 個の推論結果 $C_*^1 \sim C_*^n$ の和集合として表される．この推論結果出力について図 3.7 に示す．なお，簡単のために $n = 2$ としている．また，図の μ_{C^0} は集合 C^0 の持つメンバシップ関数である．

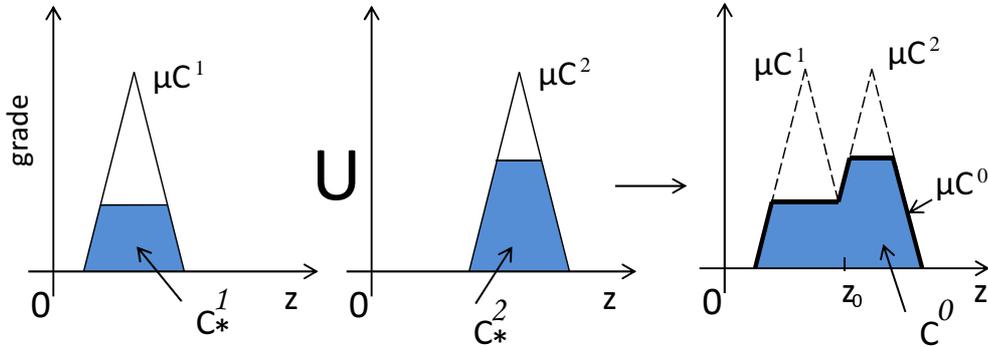


図 3.7: 複数の推論結果の統合

このような処理を行う部分は一般にファジィ推論におけるファジィ制御部と呼ばれ，規則適用の手続きを行う．

ファジィ推論では推論結果 C^0 を求めれば終了するが，本研究におけるパフォーマンス，認知負荷の推定では各指標の評価値としてファジィ集合 C^0 を求めても，実際に評価することはできない．したがって，推論結果を適切に解釈し，ある 1 つの確定した数値に変換した評価値 z^0 を決定する必要がある．ここで C^0 を次のように解釈して， z^0 に変換する．

$$z^0 = \frac{\sum_{i=1}^n z_i \cdot \mu_{C^0}(z_i)}{\sum_{i=1}^n \mu_{C^0}(z_i)} \quad (3.10)$$

このようにファジィ集合を確定値に変換する部分は，解釈部と呼ばれ，ファジィ推論結果の非ファジィ化と呼ばれる手続きを行う．式 3.10 は，ファジィ集合 C^0 のメンバシップ関数 μ_{C^0} の重心座標 z^0 を与えるものである．つまり，台集合の要素 z の重み μ_{C^0} による重み付き平均となっている．

以上，ファジィ制御部と解釈部で行った推論法は，ファジィ関係のマックスミニ合成規則による推論結果の重心をとっているため，マックスミニ合成重心法と呼ばれている．このほかにも，ファジィ集合 C^0 の解釈として図 3.6 のように C_i を ϕ_{C_i} で頭切りして C_*^i を求めるのではなく ϕ_{C_i} 倍に縮小する方式や， C^0 の非ファジィ化の方法として式 3.10 の代わりに中央値計算をしたり，最大値を与える台集合の要素を選択する高さ法など様々に提案されてきているが，これまでの数多くのファジィ推論を用いた研究

では上述のマックスミニ合成重心法が良好な結果を与えるとされている [29]。本研究でもこの手法を用いてファジィ推論を行った。

3.3.4 推論プログラムのアルゴリズム

これまでの項では本研究で行ったファジィ推論の理論を述べたが、本項ではこの理論に基づいて実装したファジィ推論プログラムのアルゴリズムについて説明する。プログラム作成には MathWorks, Inc. の MATLAB^[30] を使用した。

プログラムを図 3.8 に示すフローチャートに示す。括弧付き番号は後述のプログラム手順の番号に対応する。

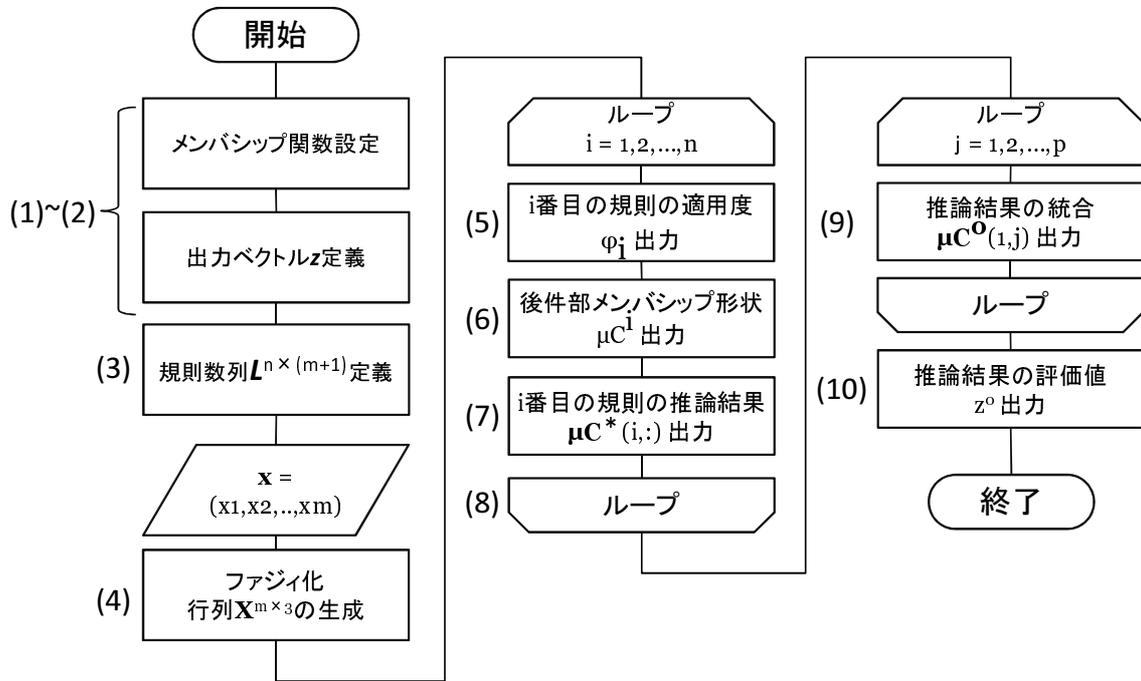


図 3.8: ファジィ推論のアルゴリズム

(1)

ファジィ推論の入力データが m 個あり、それぞれ $x_k (k = 1, 2, \dots, m)$ としたとき、

これらが各要素に格納された列ベクトル \boldsymbol{x} が与えられたとする.

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \quad (3.11)$$

(2)

この \boldsymbol{x} に対して, ファジィ推論を行い出力 z を得たい時, 前件部メンバシップ関数をそれぞれの k に対して $\mu_{Low}^k, \mu_{Middle}^k, \mu_{High}^k$ を定義する. それぞれの関数はパフォーマンス推定, 認知負荷推定, 認知・作業状態推定ごとに異なるため後述する.

一方, 後件部メンバシップ関数はパフォーマンス推定及び認知負荷推定では一律に定める. 後件部メンバシップ関数 $\mu_{Low}^0, \mu_{Middle}^0, \mu_{High}^0$ はそれぞれ

$$\mu_{Low}^0 = \begin{cases} 1 - 2|z| & (\text{if } |z| \leq \frac{1}{2}), \\ 0 & (\text{if } |z| > \frac{1}{2}), \end{cases} \quad (3.12)$$

$$\mu_{Middle}^0 = \begin{cases} 1 - 2|z - \frac{1}{2}| & (\text{if } |z - \frac{1}{2}| \leq \frac{1}{2}), \\ 0 & (\text{if } |z - \frac{1}{2}| > \frac{1}{2}), \end{cases} \quad (3.13)$$

$$\mu_{High}^0 = \begin{cases} 1 - 2|z - 1| & (\text{if } |z - 1| \leq \frac{1}{2}), \\ 0 & (\text{if } |z - 1| > \frac{1}{2}), \end{cases} \quad (3.14)$$

となる. また, 各関数の形状を図 3.9 に示す.

このような関数形状にしたのは重心法によって評価値 z^0 を算出する際に, 推論結果のファジィ集合の面積を導出するため (cf. 式 3.10), コンパクト・サポートのある関数を選択する必要があったからである. 実際, コンピュータ推論においては離散的にデータを扱うため, 厳密にはコンパクト・サポートは必要ではないが, 計算精度が落ちるため本研究では式 3.12~式 3.14 のような関数を用いた.

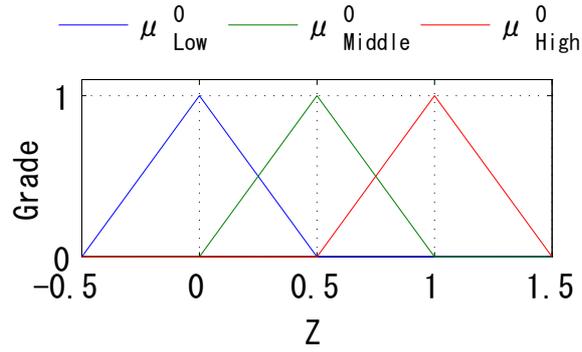


図 3.9: 後件部のメンバシップ関数形状

また出力として導出される z が取り得る範囲を $[0,1]$ とし、重心法の面積計算において z がその範囲内を全て出力し得るようにするため、式 3.12 や式 3.14 のように z が範囲 $[0,1]$ より外も取り得る関数に設定した。これにより例えば $\mu_{Low}^0 = 1$, $\mu_{Middle}^0 = 0$, $\mu_{High}^0 = 0$ であるとき $z = 0$ となる。

以上から、出力 z が取り得る値の範囲を表す以下のような行ベクトル \mathbf{z} を定義する。

$$\mathbf{z} = (z_{LB}, z_{LB} + \Delta, z_{LB} + 2\Delta, \dots, z_{UB}) \quad (3.15)$$

ここで z_{LB} 及び z_{UB} はそれぞれ z の下界と上界であり、式 3.12 や式 3.14 から $z_{LB} = -0.5$, $z_{UB} = 1.5$ である。また Δ はサンプリング間隔である。サンプリング間隔は細かい方が精度の高い推論結果が得られるが計算負荷は高くなるため、本研究では \mathbf{z} は 10^3 点程度で良いと考え、 $\Delta = (z_M - z_m)/10^3 = 2 \times 10^{-3}$ とした。

(3)

n 個の規則 $R_i (i = 1, 2, \dots, n)$ をあらかじめ定義し、規則内の集合部分について $Low = 1$, $Middle = 2$, $High = 3$ として集合番号を割り振った規則行列 $\mathbf{L}^{n \times (m+1)}$ を生成する。例えば、

R^i : If x_1 is *Low* x_2 is *Middle* \dots x_m is *High* then z is *Middle*

と定義した場合の規則行列の i 行成分 $\mathbf{L}(i, :)$ は

$$\begin{aligned} \mathbf{L}(i, [1 : m]) &= (1, 2, \dots, 3) \\ \mathbf{L}(i, m + 1) &= 2 \end{aligned} \quad (3.16)$$

となる。 n 個の規則についても前件部メンバシップ関数と同様、推定対象によって異

なるため後述する.

(4)

\mathbf{x} が入力されたとき, 各要素に対して前件部メンバシップ関数を適用して, ファジィ化処理を行いファジィ変数行列 \mathbf{X} を出力する.

$$\mathbf{X} = \begin{bmatrix} \mu_{Low}^1(x_1) & \mu_{Middle}^1(x_1) & \mu_{High}^1(x_1) \\ \mu_{Low}^2(x_2) & \mu_{Middle}^2(x_2) & \mu_{High}^2(x_2) \\ \vdots & \vdots & \vdots \\ \mu_{Low}^m(x_m) & \mu_{Middle}^m(x_m) & \mu_{High}^m(x_m) \end{bmatrix} \quad (3.17)$$

(5)

i 番目の規則に対する適応度 ϕ_i を以下のように出力する.

$$\phi_i = \min(\mathbf{X}(1, \mathbf{L}(i, 1)), \mathbf{X}(2, \mathbf{L}(i, 2)), \dots, \mathbf{X}(m, \mathbf{L}(i, m))) \quad (3.18)$$

(6)

i 番目の規則の後件部メンバシップ関数形状を離散的に行ベクトル $\mu\mathbf{C}^i$ で表現する.

$$\mu\mathbf{C}^i = \begin{cases} \mu_{Low}^0(\mathbf{z}) & (\text{if } \mathbf{L}(i, m+1) = 1), \\ \mu_{Middle}^0(\mathbf{z}) & (\text{if } \mathbf{L}(i, m+1) = 2), \\ \mu_{High}^0(\mathbf{z}) & (\text{if } \mathbf{L}(i, m+1) = 3) \end{cases} \quad (3.19)$$

(7)

図 3.6 の頭打ち集合 C_*^i のような規則 R^i の推論結果を以下のように算出し, $\mu\mathbf{C}^*$ の i 行成分へ格納する.

$$\mu\mathbf{C}^*(i, :) = \min(\mu\mathbf{C}^i, \phi_i) \quad (3.20)$$

このときの \min 関数は, 入力引数が行ベクトル $\mu\mathbf{C}^i$ 及びスカラー ϕ_i であるので, $\mu\mathbf{C}^i$ の各要素と ϕ_i とを比較し小さい方を返す関数として機能し, $\mu\mathbf{C}^i$ と同サイズの行ベクトルを返す.

(8)

$i = 1 \sim n$ で上記の(5)～(7)を繰り返す.

(9)

得られた行列 μC^* の各列 $j(j = 1, 2, \dots, p(p$ は μC^* の列数)) で, 図 3.7 に示したような推論結果の統合を行い, 全ての j で繰り返すことで, 行ベクトル μC^0 を生成する.

$$\mu C^0(1, j) = \max(\mu C^*(:, j)) \quad (3.21)$$

(10)

重心法を用いて, 推論結果の評価値 z^0 を算出する.

$$z^0 = \frac{\sum_{k=1}^p z(1, k) \cdot \mu C^0(1, k)}{\sum_{k=1}^p \mu C^0(1, k)} \quad (3.22)$$

本手法でのパフォーマンスおよび認知負荷の推定は入力引数, メンバシップ関数および規則が異なるのみであり, アルゴリズムに相違はない.

次節以降では, まず 3.4 節で執務者に行わせる認知タスクについて説明し, 3.5 節でパフォーマンス推定手法, 3.6 節で認知負荷推定手法について説明し, 最後に 3.7 節で認知・作業状態推定手法について説明する.

3.4 認知タスク

本研究では認知タスクに暗算加算タスクを用いた. 採用理由として, 暗算タスクは (1) 短期記憶および作業記憶に負荷のかかるタスクであること, (2) 難易度調整が容易であること, (3) 数字のみを使用したタスクであることの以上 3 点が挙げられる. 1 点目について, このようなタスクはこれまでの認知負荷研究でよく用いられており, それぞれの記憶にかかる負荷によって認知負荷が変動するため, これらの記憶に負荷をかけることは認知負荷計測を行う上で最も重要である. 2 点目の難易度調整の容易さについては, 難易度により変動すると考えられる執務者の内的要因 (認知負荷) を評価するため, 提案手法の妥当性評価を行う上で必要である. 最後に 3 点目については, 高度な技能を必要とするタスクではないため, 個人間の言語能力や数学能力の差によるタスクパフォーマンスへの影響が低いと考えられる.

パフォーマンス推定や認知負荷推定の評価値は知的作業によって決まるものであり、それぞれの値が知的能力を必要としない単純なキー入力の操作ミスなどから影響を受けることは最大限避けるべきである。したがって、解答入力にキーボード操作を要求した場合、キーボードの入力ミスが生じる可能性がある。そこで本研究ではPCのマウス操作のみでタスク遂行ができるよう、図3.10の(A)、(B)および(C)に示したように画面内に解答入力パネルを設置し、ボタンをキーボードのキーより大きくすることで入力ミスを抑えるよう設計した。タスクの解答手順は以下の通りである。(1)最初に図3.10(A)にあるように画面に呈示された数字を記憶した後、画面右下の「E」ボタンをクリックする。(2)次に、図3.10(B)のように呈示された同じ桁数の異なった数字と(1)で記憶した数字を足し合わせる。(3)最後に画面内の数字パネルをクリックして解答を入力し、「E」ボタンをクリックし解答を終了し、(1)へ戻る。このとき、解答入力パネルのボタンをクリックすると同時に(2)で呈示されていた数字は消え、入力した解答が図3.10(C)のように表示される。以上の(1)～(3)が1つのタスクシーケンスであり、一定時間このシーケンスを繰り返す。

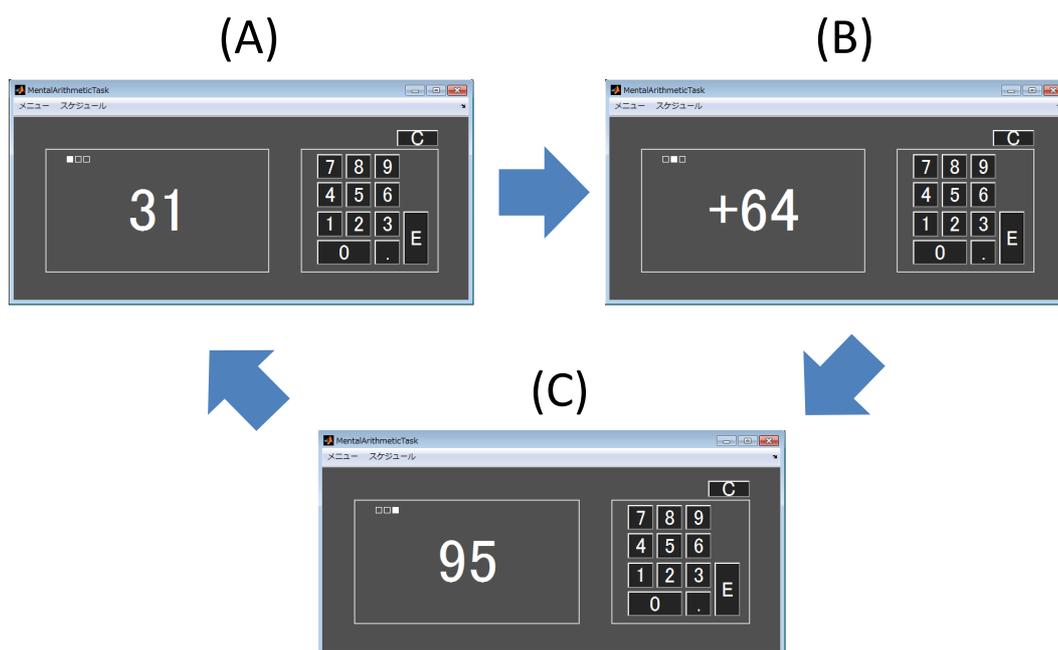


図 3.10: タスクのインタフェースとタスク解答の流れ

難易度1のタスクでは1桁の数字を提示し、難易度が上がるごとに桁数が1ずつ増加する。また、各シーケンス間での難易度を均一にするために、加算による繰り上がりが生じないように呈示数字を調整した。タスクの作業記録(タスクログ)は、(1)～(3)の作業を終えるまでのそれぞれの時間の合計(以降、解答時間と記す)および解答の正

誤である。

本研究ではパフォーマンス、認知負荷および認知・作業状態推定に複数の難易度のタスクを用いた。これは先述の通り、難易度による各指標の変動を評価するためである。したがって、少なくとも低難易度と高難易度の2群を設ける必要があり、さらにそれぞれの難易度が過度でないことが望ましい。短期記憶の観点から、暗算加算タスクで要求される最大記憶容量は、最初に覚える数字の桁数と解答する数字の桁数の和であると考えられる。よって、本研究で用いた認知タスクは同じ難易度であれば桁数が一定であるため、最大必要記憶容量は桁数×2チャンクである。人間が短期的に保持できる最大記憶容量は 7 ± 2 チャンクと言われており^[31]、高難易度では7前後のチャンクを必要とする桁数が適していると考えられる。したがって、高難易度には最大 $4 \times 2 = 8$ チャンク必要な4桁暗算を用いた。一方、低難易度タスクは、先に述べたように短期記憶容量が最大 7 ± 2 チャンクであることから、 $7 - 2 = 5$ チャンク未満を必要とするタスクが望ましい。したがって、最大 $2 \times 2 = 4$ チャンク必要とする2桁暗算もしくは $1 \times 2 = 2$ チャンク必要とする1桁暗算が該当するが、1桁暗算は暗算とは言い難く、あまりにも短期記憶に負荷がかからないと思われるので、2桁暗算を低難易度タスクとして用いた。本研究では、この2種類の難易度でそれぞれパフォーマンス、認知負荷および認知・作業状態を推定した。

3.5 パフォーマンス推定

パフォーマンスとは2.2.1項で述べた通り、作業記録から導出される作業成績のことを指すが、研究によって作業成績の定義は様々である。特に、Wargockiら^[4]は、テキストタイピングの1分あたりにタイプした文字数、単語数、および行数を作業成績としており、下田ら^[2]や岩川ら^[3]の開発したCPTOP及びCPTOP2のタスクセットは作業記録として、タスク1問あたりの解答時間や正誤を出力し、これを執務者の作業成績としている。以上のように知的生産性研究の多くは、認知タスクを執務者に行わせ、主にその解答速度やエラー率を導出することで、作業成績つまりパフォーマンスを評価している。また、宮城ら^[6]は執務者は常に作業を行っているのではなく、マルコフモデルに基づいた複数の作業状態遷移をしながら作業をしており、その遷移確率は疲労やモチベーションなどの内的要因によって変化すると示唆している。さらに、タスク1問あたりの解答時間のヒストグラムが複数の対数正規分布に近似されることを経験的に導いている。したがって、執務者のタスク解答時間は必ずしも一定ではなく分

散し、その分散値もパフォーマンス評価に重要な因子であると考えられる。

以上より、本研究ではパフォーマンスとは (1) タスク解答時間などの作業時間、(2) 解答時間の分散で表される作業時間の分散性、(3) 作業のエラー率という、3つの特徴量で構成される指標、言わば作業の「質」とであると定義した。以降ではそれぞれを解答時間、分散性、エラー率と記す。つまり、最もパフォーマンスが高い状態とは、解答時間が短く、かつ分散性が低く、さらにエラー率が低い状態である。この定義の下、本研究ではパフォーマンス推定のための入力引数と規則を設定した。つまり、3.3.3項において x_1 を解答時間、 x_2 を分散性、 x_3 をエラー率とし、 z をパフォーマンスとして推定した。

3.4項で認知タスクについて述べた後、その次項以降は特徴量抽出やファジィ推論におけるメンバシップ関数および規則の設定について述べる。

3.5.1 特徴量抽出

n 個の解答タスクログがあり、解答タスク番号 i ($i = 1, 2, \dots, n$) についてそれぞれ解答時間 ST_i 、正誤 RW_i (正: 1, 誤: 0) が与えられたとき、エラー率 E_i および分散性 V_i を、 $i - (p - 1)$ 番目から i 番目の計 p 点分のタスクの解答時間や正誤のデータから導出した。 $i \geq p$ のとき、 E_i および V_i はそれぞれ式 3.23、式 3.24 で表される。

$$E_i = 1 - \frac{1}{p} \sum_{k=i-(p-1)}^i RW_k \quad (3.23)$$

$$V_i = \frac{1}{p-1} \sum_{k=i-(p-1)}^i (ST_k - \overline{ST}^*)^2 \quad (3.24)$$

$$\overline{ST}^* = \frac{1}{p-1} \sum_{k=i-(p-1)}^i ST_k \quad (3.25)$$

このとき、解析窓サイズ p を大きく設定すると、特徴量抽出に要する時間が増加し、時間分解能が低下する恐れがある。逆に p を小さくとると抽出が正しく行われな可能性がある。これまで、妥当な解析窓サイズ p を明確に示唆した研究や知見はないが、パフォーマンスが1分前後で急激に変化するとは考えにくいいため、本研究では解析窓は最長1分前後が妥当であると考え、タスク1問の解答に最大10秒以上を必要とするとして仮定してタスク5問を解析窓として設定、つまり $p = 5$ とした。以上のように、特徴量は $i = p$ 番目以降のタスクログから出力した。

3.5.2 メンバシップ関数の設定

各特微量に対するメンバシップ関数(前件部メンバシップ関数)を式3.26～式3.31に示した.

$$\mu_{Low}(x) = \begin{cases} 1 - \exp\left\{-\frac{(x - T_{High})^2}{2\sigma_{Low}^2}\right\} & (\text{if } x \leq T_{High}), \\ 0 & (\text{if } x > T_{High}), \end{cases} \quad (3.26)$$

$$\sigma_{Low} = \frac{T_{High} - T_{Low}}{h} \quad (3.27)$$

$$\mu_{Middle}(x) = \exp\left\{-\frac{(x - T_{Middle})^2}{2\sigma_{Middle}^2}\right\} \quad (3.28)$$

$$\sigma_{Middle} = \begin{cases} \frac{T_{Middle} - T_{Low}}{h} & (\text{if } x \leq T_{Middle}), \\ \frac{T_{High} - T_{Middle}}{h} & (\text{if } x > T_{Middle}), \end{cases} \quad (3.29)$$

$$\mu_{High}(x) = \begin{cases} 1 - \exp\left\{-\frac{(x - T_{Low})^2}{2\sigma_{High}^2}\right\} & (\text{if } x \geq T_{Low}), \\ 0 & (\text{if } x < T_{Low}) \end{cases} \quad (3.30)$$

$$\sigma_{High} = \frac{T_{High} - T_{Low}}{h} \quad (3.31)$$

ここで, x は入力データであり, 解答時間 ST , 分散性 V もしくはエラー率 E のいずれかである. h は, 解答時間および分散性では2, エラー率では3とした. また, T_{Low} , T_{Middle} 及び T_{High} はメンバシップ関数の Low, Middle および High それぞれの閾値であり, 各特微量別に異なった値を設定した. σ_{Low} , σ_{Middle} 及び σ_{High} はガウシアン型メンバシップ関数の幅である. メンバシップ関数の形状を図3.11に示す.

μ_{Low} および μ_{High} の関数形状は, それぞれ式3.5のZ型関数および式3.4のS型関数をガウシアン形状にしたものであり, μ_{Middle} は式3.7の形状と同じである. ガウシアン形状の関数を用いた理由として, 3.3.2項で述べた通り, 各特微量のメンバシップ関

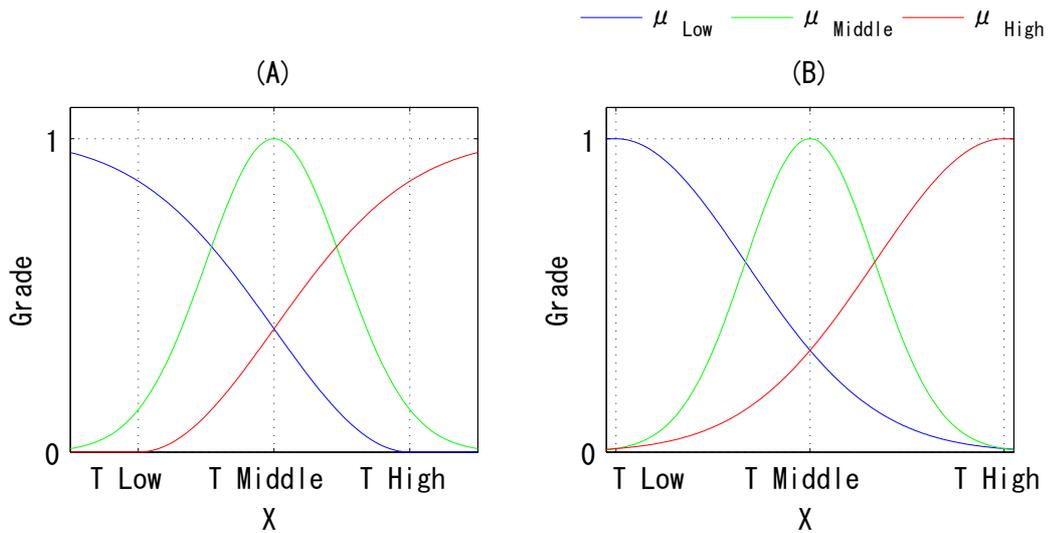


図 3.11: 入力特徴量に対するメンバシップ関数の形状

数の閾値について明確に示唆された既往研究が少ないことが挙げられる．本研究では入力データを閾値前後でファジィ化するとき，コンパクト・サポートのないガウシアン形状の関数を用いることで，ファジィ変数がすぐに0もしくは1に収束しないよう，閾値をある程度超えるもしくは下回ることを許容した．しかし，どの程度許容するかについての議論は困難である．ここで，正規分布の平均値を m ，標準偏差 σ とした場合， $m \pm \sigma$ ， $m \pm 2\sigma$ ， $m \pm 3\sigma$ はそれぞれ正規分布の面積の 68.26%，95.44%，99.72% を囲む値である． $m \pm \sigma$ が閾値になるようメンバシップ関数を設定した場合，外側面積で約 30% 許容することになり，これは許容範囲が大きい． $m \pm 3\sigma$ はほぼ 100% に近く，許容範囲が小さい．したがって，解答時間や分散性のメンバシップ関数では， $m \pm 2\sigma$ が最も適していると考え，閾値 T_{Low} ， T_{High} が図 3.11(A) に示すようにガウシアン分布の 95% を含む程度になるようメンバシップ関数を設定した．したがって， σ_{Low} について， $m = T_{Low}$ とし， $m + 2\sigma = T_{High}$ とした場合， $\sigma_{Low} = (T_{High} - T_{Low})/2$ となり，式 3.27 の h は 2 となる．同様に σ_{Middle} や σ_{High} についても計算すると，式 3.29 や式 3.31 でも $h = 2$ となる．一方，エラー率の閾値は 0% や 100% などある程度明確に決められるため，やや議論は残るものの $m \pm 3\sigma$ が適していると考え，図 3.11(B) のように設定した．同様に計算すると $h = 3$ となる．

次に，前件部メンバシップ関数のパラメータ T_{Low} ， T_{Middle} ， T_{High} について述べる．以降では，区別のために解答時間におけるメンバシップのパラメータを T_{Low}^{ST} ， T_{Middle}^{ST} ， T_{High}^{ST} ，分散性におけるパラメータを T_{Low}^V ， T_{Middle}^V ， T_{High}^V ，エラー率におけるパラメー

タを T_{Low}^E , T_{Middle}^E , T_{High}^E と記す. 本研究ではこれらのパラメータ設定のために, まず解答時間の理論値を導出しなければならない. 暗算加算タスクは数字情報の記憶・想起・演算が知的処理の中心であり, Card ら^[32]の提案する人間情報処理モデルにより本タスクの解答時間の理論値が計算できる. 河野による知的生産性のシミュレーション^[33]では, この理論値が実測値とよく整合することが示されている. そこで, 本研究では, 2桁暗算や4桁暗算をそれぞれ表3.1や表3.2のように認知ステップ分解し, 各ステップの処理時間の最短と最長の合計から解答時間の理論値を導出した. 以上の計算から, 各難易度の解答時間の理論値を表3.3に示した.

この解答時間の理論値は, 認知タスクを解答するための認知ステップが必要最小限のステップ数でかつ, それぞれが順次処理されることを想定して導出された値である. このような想定をしたのは, そもそもこの理論値はメンバシップ関数のパラメータに使用するための値であり, そのパラメータが過度な値になることを避ける必要があるためである. 例えばタスクへの習熟による認知ステップの並列処理や, 慎重に解答するために計算を複数回行うなどの新たな認知ステップの追加という特殊な場合は非常に多様に存在するため, それらを理論値導出で考慮した場合, 極端に短いもしくは長い解答時間の理論値が導出される可能性がある. そうなれば, 解答時間の変動に対してパフォーマンス推定の感度が下がり, 急激に解答時間が変化しなければパフォーマンスの推定値がほとんど変化せず, パフォーマンス推論器の意義が失われる.

したがって, 本研究ではパラメータ設定に用いる解答時間の理論値は, 必要最低限の認知ステップが順次処理される時の解答時間と設定した. 実際に執務者が並列処理によって解答時間を短縮し, 解答時間の理論値より短い時間でタスクを遂行した場合は, 本手法によるとパフォーマンスの推定値が1に近い値をとる可能性が高くなるが, これはタスクへの習熟が進み, 結果として高いパフォーマンスが発揮されていると解釈できる. 一方, 認知ステップが追加されることで解答時間が延長され, 解答時間の理論値より長くなった場合は, パフォーマンスの推定値は0に近い値をとる可能性があるが, このときは作業への集中が乱れ, 本来必要のない認知ステップを処理しており, パフォーマンスが低くなったと解釈できる.

表 3.1: 2桁暗算タスクの認知ステップと最短および最長処理時間

順番	ステップ	分類	最短 (ms.)	最長 (ms.)
1	十の位の数字を見ようとする	運動命令	30	100
2	十の位の数字を見る	眼球運動	20	50
3	十の位の数字を確認	知覚	50	200
4	十の位の数字を記憶	記憶	25	25
5	一の位で, 順番 1~4 と同様	上記と同様	125	375
6	Enter キーを押す	運動	500	900
7	十の位の数字を見ようとする	運動命令	30	100
8	十の位の数字を見る	眼球運動	20	50
9	十の位の数字を確認	知覚	50	200
10	十の位の数字を記憶	記憶	25	25
11	最初の十の位の数字を思い出す	想起	175	175
12	十の位の数字を足し合わせる	思考	25	170
13	十の位の数字を記憶	記憶	25	25
14	一の位で, 順番 10~13 と同様	上記と同様	350	745
15	保持情報を整理する	想起	175	175
16	十の位を思い出す	想起	175	175
17	キーを押す	運動	500	900
18	一の位を思い出す	想起	175	175
19	キーを押す	運動	500	900
20	Enter キーを押す	運動	500	900
		計	3475	6365

表 3.2: 4桁暗算タスクの認知ステップと最短および最長処理時間

順番	ステップ	分類	最短 (ms.)	最長 (ms.)
1	千の位の数字を見ようとする	運動命令	30	100
2	千の位の数字を見る	眼球運動	20	50
3	千の位の数字を確認	知覚	50	200
4	千の位の数字を記憶	記憶	25	25
5	百の位で, 順番 1~4 と同様	上記と同様	125	375
6	十の位で, 順番 1~4 と同様	上記と同様	125	375
7	一の位で, 順番 1~4 と同様	上記と同様	125	375
8	Enter キーを押す	運動	500	900
9	千の位の数字を見ようとする	運動命令	30	100
10	千の位の数字を見る	眼球運動	20	50
11	千の位の数字を確認	知覚	50	200
12	千の位の数字を記憶	記憶	25	25
13	最初の千の位の数字を思い出す	想起	175	175
14	千の位の数字を足し合わせる	思考	25	170
15	千の位の数字を記憶	記憶	25	25
16	百の位で, 順番 10~15 と同様	上記と同様	350	745
17	十の位で, 順番 10~15 と同様	上記と同様	350	745
18	一の位で, 順番 10~15 と同様	上記と同様	350	745
19	保持情報を整理する	想起	175	175
20	千の位を思い出す	想起	175	175
21	キーを押す	運動	500	900
22	百の位を思い出す	想起	175	175
23	キーを押す	運動	500	900
24	十の位を思い出す	想起	175	175
25	キーを押す	運動	500	900
26	一の位を思い出す	想起	175	175
27	キーを押す	運動	500	900
28	Enter キーを押す	運動	500	900
		計	5745	10655

表 3.3: 解答時間の理論値

桁数	最小値 (sec.)	最大値 (sec.)
2 桁	3.375	6.165
4 桁	5.645	10.455

解答時間

執務者は、表 3.1 や表 3.2 の認知ステップを経てタスクに解答していると仮定し、 T_{Low}^{ST} および T_{High}^{ST} を表 3.3 における解答時間の理論値の最小値、最大値とした。また、 T_{Middle}^{ST} は T_{Low}^{ST} と T_{High}^{ST} の相加平均とした。

分散性

T_{Low}^V を分散がない場合の閾値と仮定し、0 とした。次に、 T_{Middle}^V を T_{Low}^{ST} 付近及び T_{Middle}^{ST} 付近の解答時間 ST が交互に計測されている時の分散値、つまり理論上は中程度に解答時間にバラつきがある時の分散値であると仮定し、

$$\begin{aligned}
 T_{Middle}^V &= \left(\frac{T_{Middle}^{ST} - T_{Low}^{ST}}{2} \right)^2 \\
 &= \left(\frac{T_{High}^{ST} - T_{Low}^{ST}}{4} \right)^2 \tag{3.32} \\
 (\because T_{Middle}^{ST} &= \frac{T_{High}^{ST} + T_{Low}^{ST}}{2})
 \end{aligned}$$

と設定した。

これは、 T_{Middle}^V を T_{High}^{ST} 付近及び T_{Middle}^{ST} 付近の ST が交互に計測されている時の分散値と仮定しても同じ値が導出される。

同様に T_{High}^V は理論上最も解答時間にバラつきがある時の分散値、つまり T_{High}^{ST} 付近及び T_{Low}^{ST} 付近の ST が交互に計測されている時の分散値と仮定した。

$$T_{High}^V = \left(\frac{T_{High}^{ST} - T_{Low}^{ST}}{2} \right)^2 \tag{3.33}$$

エラー率

エラー率は、最小値が 0% 最大値が 100% であるので、 $T_{Low}^E = 0\%$ 、 $T_{Middle}^E = 50\%$ 、 $T_{High}^E = 100\%$ とした。

3.5.3 規則設定

パフォーマンス推定では表 3.4 に示す規則で推論を行った。解答時間、分散性及びエラー率の 3 変数に対してそれぞれ「High」「Middle」「Low」の 3 ファジィ集合を設定したため、規則数は $3 \times 3 \times 3 = 27$ 個となった。例えば No.1 の場合、「If 解答時間 is High and 分散性 is High and エラー率 is High then パフォーマンス is Low」というように規則が設定されているとなる。

パフォーマンスは解答時間、分散性及びエラー率の 3 要素によって決定され、それぞれが低ければパフォーマンスは良く、逆にそれぞれが高い場合はパフォーマンスが低い。また、解答時間や分散性が低い場合でも、エラー率が中程度もしくは高い場合は意図的なエラーや勘違いによるエラーが考えられるためパフォーマンスは低いが、エラー率が低い場合はパフォーマンスは高いと言える。逆に、エラー率が低い場合でも解答時間や分散性が高い時は、正確性を求めるあまり作業量が低減していると考えられるためパフォーマンスは低いと考えられる。以上のような考えに基づき設定した規則を表 3.4 に示す。

表 3.4: パフォーマンス推定での規則

No.	前件部			後件部
	解答時間	分散性	エラー率	パフォーマンス
1	High	High	High	Low
2	High	High	Middle	Low
3	High	High	Low	Low
4	High	Middle	High	Low
5	High	Middle	Middle	Low
6	High	Middle	Low	Middle
7	High	Low	High	Low
8	High	Low	Middle	Low
9	High	Low	Low	Middle
10	Middle	High	High	Low
11	Middle	High	Middle	Middle
12	Middle	High	Low	Middle
13	Middle	Middle	High	Low
14	Middle	Middle	Middle	Middle
15	Middle	Middle	Low	High
16	Middle	Low	High	Low
17	Middle	Low	Middle	Middle
18	Middle	Low	Low	High
19	Low	High	High	Low
20	Low	High	Middle	Middle
21	Low	High	Low	Middle
22	Low	Middle	High	Low
23	Low	Middle	Middle	Middle
24	Low	Middle	Low	High
25	Low	Low	High	Low
26	Low	Low	Middle	High
27	Low	Low	Low	High

3.6 認知負荷推定

認知負荷はアンケートなどの主観指標や、タスクの作業記録からの行動指標もしくは執務者自身から計測可能な生理指標などの客観指標から測定できるとされている^[7]。主観指標は執務者の自己申告ではあるが、認知負荷を測定するには有効な指標である。しかし、時間分解能は低く、分解能向上のためにアンケートなどの回数を増やすと作業の阻害を招き、それ自体が認知負荷を増加させる恐れがあるため、リアルタイム測定は困難である。一方、行動指標として、入力ペンの筆跡の変化から認知負荷を計測する研究^[11]や、音声タスクを用いた研究^[12]があるが、タスクから計測できる認知負荷が真に執務者にかかる負荷であるという根拠は現在も乏しく、行動指標による認知負荷測定の信頼性は十分でない。一方、Fredら^[8]は、認知負荷を執務者の生理指標から測定することの妥当性を示唆しており、生理学分野では認知負荷と生理指標の密接な関係が多くの研究で証明されている。特に、瞳孔径^[17-21]、心拍^[22-27]は認知負荷をよく反映する指標として有効であると示唆されており、これら2指標を用いることで執務者の実際の認知負荷を計測できると考えられる。さらに、近年の技術発展により、非侵襲のみならず非接触の測定デバイスが登場しており^[34-36]、将来的には完全非接触の認知負荷計測も期待される。

以上の理由から本研究では、認知負荷推定のためのファジィ推論の入力データに瞳孔径および心拍数を採用した。つまり、3.3.3項において x_1 を瞳孔径、 x_2 を心拍数とし、 z を認知負荷として推定を行った。

3.6.1項では各指標からの特徴量抽出の詳細を述べ、3.6.2項、3.6.3項ではそれぞれメンバシップ関数および規則設定について述べる。

3.6.1 特徴量抽出

本項では、先に瞳孔径について瞬目によるアーチファクト除去と特徴量抽出を述べ、次に心拍数の導出法と特徴量抽出について述べる。

瞳孔径

瞳孔径計測には多くの計測手法が存在するが、一般的な手法として、赤外線カメラなどで執務者の顔を撮影し、顔認識システムから眼球位置を自動認識した後、瞳孔径を計測する Camera Eye-tracking システムが挙げられる (faceLAB^[36])。近年の高度な Eye-tracking システムでは、サンプリング周波数が 60Hz 以上と高い。そのため、瞳孔径だけでなく瞬目や閉瞼率も精度よく計測できるが、同時に瞬目によって瞳孔径が瞼

に隠れる時も計測されるため、瞳孔径のトラッキングエラーを起こすことがある。したがって、計測された瞳孔径をそのまま用いるのは問題がある。トラッキングエラー時の瞳孔径がどのように評価されるかは Eye-tracking システムの仕様によるが、多くの場合、非常に小さい値もしくは大きい値が評価値として出力される。本研究で使用する faceLAB では、瞼が瞳孔径の数% 重なる程度では精度の低下は起こらないが、瞼が瞳孔径の大半を覆ってしまうと精度は急激に落ち、過度に大きな評価値が返されることが多くなる。さらに、完全に瞼が瞳孔径を覆った場合は、0 を返すようになる。本研究では、この瞬目による瞳孔径データのアーチファクトを除去するために、(1) 一定の閾値を超えるデータと (2) 急激な評価値の変動 (データジャンプ) を消去した。(1) の除去方法については、薬物や電気刺激を用いない一般的な場合の健常者の瞳孔径は約 2mm(明所) から約 8mm(暗所) まで変化する^[37] ため、2mm を下回るもしくは 8mm を上回る瞳孔径データを消去した。また、(2) については Marchall^[38] が瞳孔径データジャンプの除去方法を提案しており、サンプリング周波数が 60Hz のトラッキングシステムによる瞳孔径データでは、2 点前 (1/30 秒前) の計測データと比べて 0.1mm 以上異なるデータ点はアーチファクトとして消去すべきであると示唆している。本研究でも瞳孔径計測にはサンプリング周波数が 60Hz の計測カメラを用いたため、この除去方法を採用した。以上 (1) 及び (2) の手法で瞬目によるアーチファクトを除去した例を図 3.12 に示す。

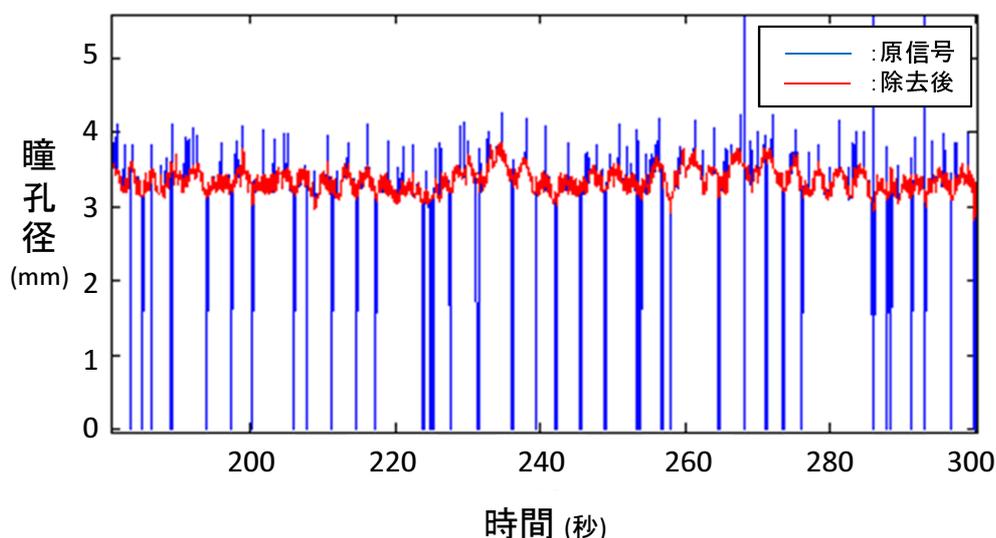


図 3.12: 瞬目による瞳孔径アーチファクトの除去

青線が faceLAB から出力された瞳孔径データであり、その線に重畳して除去後のデータを赤線で示した。図 3.12 より、瞬目による不自然なデータの降下やジャンプが適切

に除去されていることが分かる。

アーチファクト除去の後，瞳孔径データから認知負荷に関する特徴量を抽出した。瞳孔径は認知負荷によって散大し，負荷が解消された後，収縮が始まりピークが形成される^[17]。つまり，瞳孔径のピークが認知負荷量に関係すると考え，本研究では図 3.13 に示すように， $t_{i-1} \sim t_i$ の区間から t' だけ遅れた区間を抽出区間 i とし，その区間内の瞳孔径ピーク値(極大値)を，タスク遂行にかかる認知負荷を反映する特徴量 PD_i として特徴量抽出を行った。

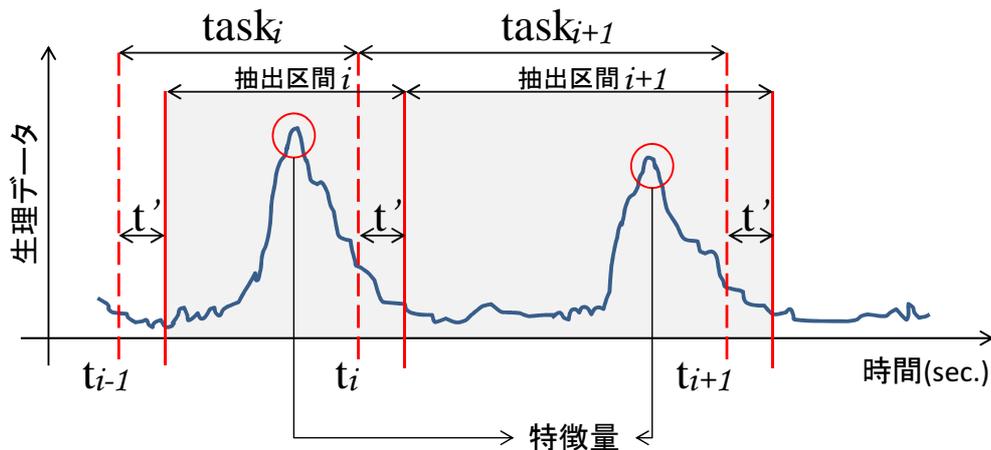


図 3.13: タスクごとの特徴量抽出

図中の t_i は最初のタスクを開始した時刻を $t = 0$ としたときの， i 番目のタスク解答が終了した時刻であり，3.5.1 項での解答時間 ST_i を用いて，

$$t_i = \sum_{k=1}^i ST_k \quad (3.34)$$

と表せる。また，3.5.1 項より，パフォーマンス推定の特徴量出力を 5 番目のタスクログから行っているため，瞳孔径の特徴量においても $i \geq 5$ で抽出を行った。さらに，抽出区間 i を t_{i-1} から t_i の間ではなくその区間から t' だけ遅れた区間としたのは，瞳孔径のピーク反応はタスク呈示からある潜時をおいて現れるため，その反応の遅れに合わせて抽出区間を調整する必要があったからである。既往研究^[39]によると，被験者が暗算積算している時の瞳孔径を計測した実験では，タスクが呈示されてから約 2 秒から 5 秒の間でピークが観測されているため，タスク呈示の直後から認知負荷がかかっていると仮定すると，認知負荷を表すピークは認知負荷がかかり始めてから最大で 5 秒程度遅れた時に観測されると考えられる。したがって，本研究では $t' = 5(sec.)$ として特

微量を抽出した。

心拍数

心拍数は一般に心電図 (Electrocardiogram, ECG) による心筋の筋電位計測から評価することが多い。昔からよく知られた方法であり、現在でも生理学研究や医療の分野で用いられている。また、近年では微弱なマイクロウェーブを体に照射し、反射マイクロウェーブから心拍時の微小な血圧上昇を計測することで心拍数を推定する手法^[34]や、心拍時の微小な血圧変動による顔の表面皮膚の光反射率の違いをカメラで捉え心拍数を推定する手法^[35]が開発されており、非接触の心拍計測技術が次々と提案されている。特に後者の手法はすでにスマートフォンのアプリケーションとして開発されているため、次世代の心拍計測手法として期待されている。しかし、未だ計測精度には議論が残り、研究開発段階であるため、本研究では ECG を用いて心拍数を評価した。電極を心臓に関して左右上下対称に 2 点、例えば右首筋と左脇腹にそれぞれ 1 点ずつ装着した場合、図 3.14 のような波形が計測される。

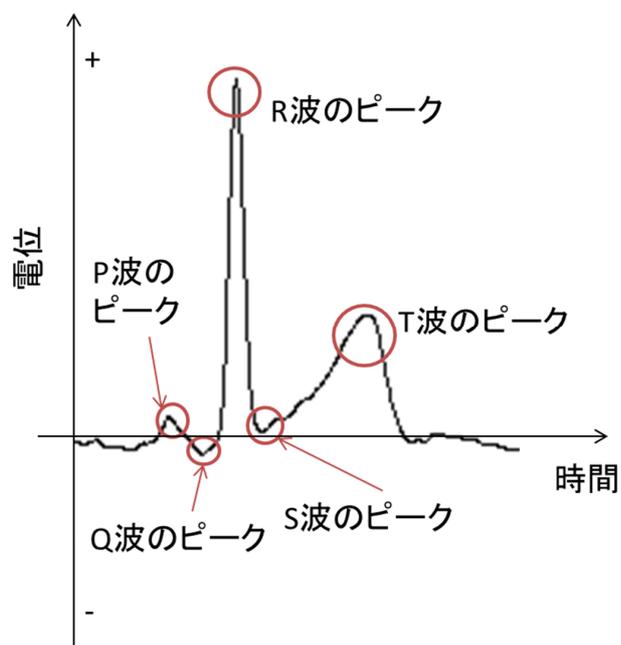


図 3.14: 心電図 (ECG) の例

心臓では洞結節と呼ばれるペースメーカの役割を果たす部位から周期的に電氣的刺激が発生し、この刺激が心臓全体に広がり、図のような P 波、QRS 波 (単に R 波とも呼ぶ)、T 波と呼ばれる一連の波形を示す。まず、洞結節による刺激が心房の脱分極 (興奮) を起こし、これが ECG 上では P 波として観測される。次に刺激が心房から心室へ移り、PQ 間の時間を経て心室の脱分極 (興奮) が起こり、心室収縮によって全身に血液

が送り出される。この時 R 波が観測される。最後に ST 間を経て心室の再分極が起こる。これは収縮した心臓を弛緩させるためである。この時 T 波が観測される。

以上が心電図から見た心臓の一連の動きであるが、一般に心拍と呼ばれるのは ECG では R 波の立ち上がりを指す。心拍数は R 波ピークの時間間隔 (秒) を 1 拍ごとに導出し、このピーク間隔を RRI (R-R interval) としたとき、 $60/RRI$ (bpm) として算出できる。人間の心拍は一般に新生児から成人になるにつれて拍数が減少していき、健全な成人では安静時に 53bpm から 95bpm の心拍数が観測される^[40]。

本研究でも ECG から R 波のピークを検出して心拍数を算出した。しかし、このままでは心拍数データは等間隔ではない時系列データとなるので、心拍変動に関する研究でよく用いられている、サンプリング周波数 10Hz での 3 次 spline 補間を行い、等間隔の時系列データに変換した。

心拍数の算出後、認知負荷に関する特徴量を抽出した。瞳孔径と同様、心拍数は認知負荷の上昇によって酸素消費量と共に上昇し、負荷が解消された後は減少しピークが生じる^[22]。よって、瞳孔径と同様に抽出区間 i の心拍数のピークを特徴量 HR_i として図 3.13 のように抽出した。また、心拍数は刺激呈示から 1 拍程度 (成人で約 0.6 ~ 1.1 秒) 遅れることが報告されており^[41]、本研究では 2 拍遅れる可能性も考慮し、やや長めに見積もって $t' = 2(sec.)$ として特徴量を抽出した。

3.6.2 メンバシップ関数の設定

各特徴量に対するメンバシップ関数 (前件部メンバシップ関数) は 3.5.2 項と同様、式 3.26~式 3.31 とし、 $h = 2$ とした。理由も同様であり、瞳孔径や心拍数もメンバシップ関数の閾値について明確に示唆された既往研究が少ないため、コンパクト・サポートのないガウシアン形状の関数を用いた。

次に、前件部メンバシップ関数のパラメータ T_{Low} , T_{Middle} , T_{High} について述べる。以降では、区別のために瞳孔径におけるメンバシップのパラメータを T_{Low}^{PU} , T_{Middle}^{PU} , T_{High}^{PU} 、心拍におけるパラメータを T_{Low}^{HR} , T_{Middle}^{HR} , T_{High}^{HR} と記した。

まず、執務者を一定時間安静にさせ、安静時は認知負荷が全くかかっていない状態であると仮定し、この時の瞳孔径及び心拍数の平均値をそれぞれ T_{Low}^{PU} , T_{Low}^{HR} に設定した。安静期間の時間設定において明確な知見はないが、極端に短い場合は安静時の正確な瞳孔径や心拍が正しく計測できず、極端に長い場合は執務者の眠気が原因で瞳孔径の収縮が進み、眠気によるアーチファクトがデータに混入する可能性がある。本研究では眠気の促進を抑えつつ正確な計測を目指すため、1 分間を安静時間とした。

瞳孔径について、Hess^[17]らによるとタスク遂行時の瞳孔径と積算の暗算タスクの難易度には相関が認められ、タスク開始前の安静時と比べて4%～29.5%程度の上昇が認められている。Hessらの実験で用いられた2桁×2桁の暗算積算は短期記憶を7チャンク前後要求するタスクであり、人間の短期記憶が最大で7±2チャンクであることを考えると、このタスクによる認知負荷は高いと言える。したがって、本研究では認知負荷が高い時、瞳孔径は安静時と比べて最大約30%散瞳すると考え、 $T_{High}^{PU} = T_{Low}^{PU} \times 1.3$ とし、 $T_{Middle}^{PU} = T_{Low}^{PU} \times 1.15$ と設定した。一方、心拍数については石橋ら^[22]によると数字の暗唱タスクにおいて心拍数と難易度に相関が認められ、特に7桁の暗唱では20%の上昇が認められた。こちらのタスクも短期記憶の観点から高い認知負荷を与えると考えられる。したがって、高い認知負荷下では心拍数は安静時に比べて約20%上昇すると考え、 $T_{High}^{PU} = T_{Low}^{PU} \times 1.2$ 、 $T_{Middle}^{PU} = T_{Low}^{PU} \times 1.1$ と設定した。

3.6.3 規則設定

認知負荷推定で設定した規則は表3.5に示す通りである。瞳孔径および心拍数の2変数に対してそれぞれ「High」「Middle」「Low」の3ファジィ集合を設定したため、規則数は $3 \times 3 = 9$ 個となった。

表 3.5: 認知負荷推定での規則

No.	If		Then
	瞳孔径	心拍数	認知負荷
1	High	High	High
2	High	Middle	High
3	High	Low	Middle
4	Middle	High	High
5	Middle	Middle	Middle
6	Middle	Low	Low
7	Low	High	Middle
8	Low	Middle	Low
9	Low	Low	Low

認知負荷推定における規則は、パフォーマンス推定での規則と比べ設定が容易である。既往研究より瞳孔径や心拍数と認知負荷の関係が認められているため、瞳孔径及び

心拍数が高ければ認知負荷は高く、逆にそれぞれ低ければ認知負荷も低いとする単純な規則設定で十分である。また、一方が高く他方が低いというほぼ逆の反応を示す場合は、過去の研究から考えると起こりにくいケースではあるが、そのような場合の認知負荷は上昇中もしくは下降中であると考え、中位程度の負荷であると仮定し、「Middle」として設定した。

3.7 認知・作業状態推定

パフォーマンス及び認知負荷を推定した後、最後に認知・作業状態を推定する。ここで、本研究で提案した図 3.1 のパフォーマンス-認知モデルをファジィ推論の規則表のようにしてまとめると、表 3.6 のようになる。

表 3.6: 認知状態推定における規則

No.	If		Then
	パフォーマンス	認知負荷	認知・作業状態
1	High	High	High Working State
2	High	Low	Optimal Working State
3	Low	High	Bad Working State
4	Low	Low	Low Working State

ここで、パフォーマンス及び認知負荷をそれぞれファジィ化し、各規則つまり各認知・作業状態の適応度を算出した。以下にその詳細を述べる。

パフォーマンスおよび認知負荷のファジィ化にあたって、設定した規則からファジィ集合は「Low」および「High」の2集合とした。また、使用したメンバシップ関数は式 3.26 および式 3.30, $h = 3$ とし、それぞれの閾値については $T_{Low} = 0$, $T_{High} = 1$ と設定した。これはパフォーマンス推定のエラー率の時と同様のパラメータであり、これは 3.5 節及び 3.6 節で、パフォーマンスと認知負荷がファジィ推論により $[0,1]$ の範囲で定量化されているためである。

ここで、ある時点での表 3.6 の各規則の適応度について High Working State の適応度を ϕ_{CH} , Optimal Working State を ϕ_{CO} , Bad Working State を ϕ_{CB} , Low Working State を ϕ_{CL} とする。さらに、このときのパフォーマンスを X_i , 認知負荷を Y_i としたとき、式 3.26~式 3.31 から算出された $\mu_{Low}(X_i)$ などのファジィ変数を用いて、適応度 ϕ_C は以下のようなになる。

$$\phi_{C_H} = \min(\mu_{High}(X_i), \mu_{High}(Y_i)) \quad (3.35)$$

$$\phi_{C_O} = \min(\mu_{High}(X_i), \mu_{Low}(Y_i)) \quad (3.36)$$

$$\phi_{C_B} = \min(\mu_{Low}(X_i), \mu_{High}(Y_i)) \quad (3.37)$$

$$\phi_{C_L} = \min(\mu_{Low}(X_i), \mu_{Low}(Y_i)) \quad (3.38)$$

以上の適応度をそれぞれの認知・作業状態の適応度として本手法の出力とした。それぞれ $[0,1]$ をとる値であり、ある状態の適応度が 0 に近いほどその状態である可能性は低く、1 に近いほどその状態である可能性が高いと言える。

執務者の認知・作業状態を一意に推定する場合は、各 4 状態の適応度を算出するだけでなく、それらの中で最も適応度の高い状態をその時点の状態と推定する。しかし、そのような推定手法の欠点として、適応度が非常に近い値を取る複数の状態が存在する場合、僅かな誤差により推定結果が大きく変わる可能性がある。一方、一意に状態を推定せず、各状態の適応度を推定結果として出力した場合、それぞれの適応度が拮抗する場合や逆にある 1 状態の適応度が明らかに高い場合などが判別可能となるため、認知・作業状態の誤推定の回避が容易になる。したがって、本研究ではあえて認知・作業状態を一意に推定せず、4 状態の適応度をそれぞれ推定結果として出力した。

最後に、3.5 節で述べたパフォーマンス推定、3.6 節で述べた認知負荷推定、そして本節で述べた認知・作業状態推定をまとめたアルゴリズムのフローチャートを図 3.15 のようなフローチャートに示す。図中のサブルーチン処理「ファジィ推論」は 3.3.4 項で定義した処理であり、フローチャートは図 3.8 に示した通りである。

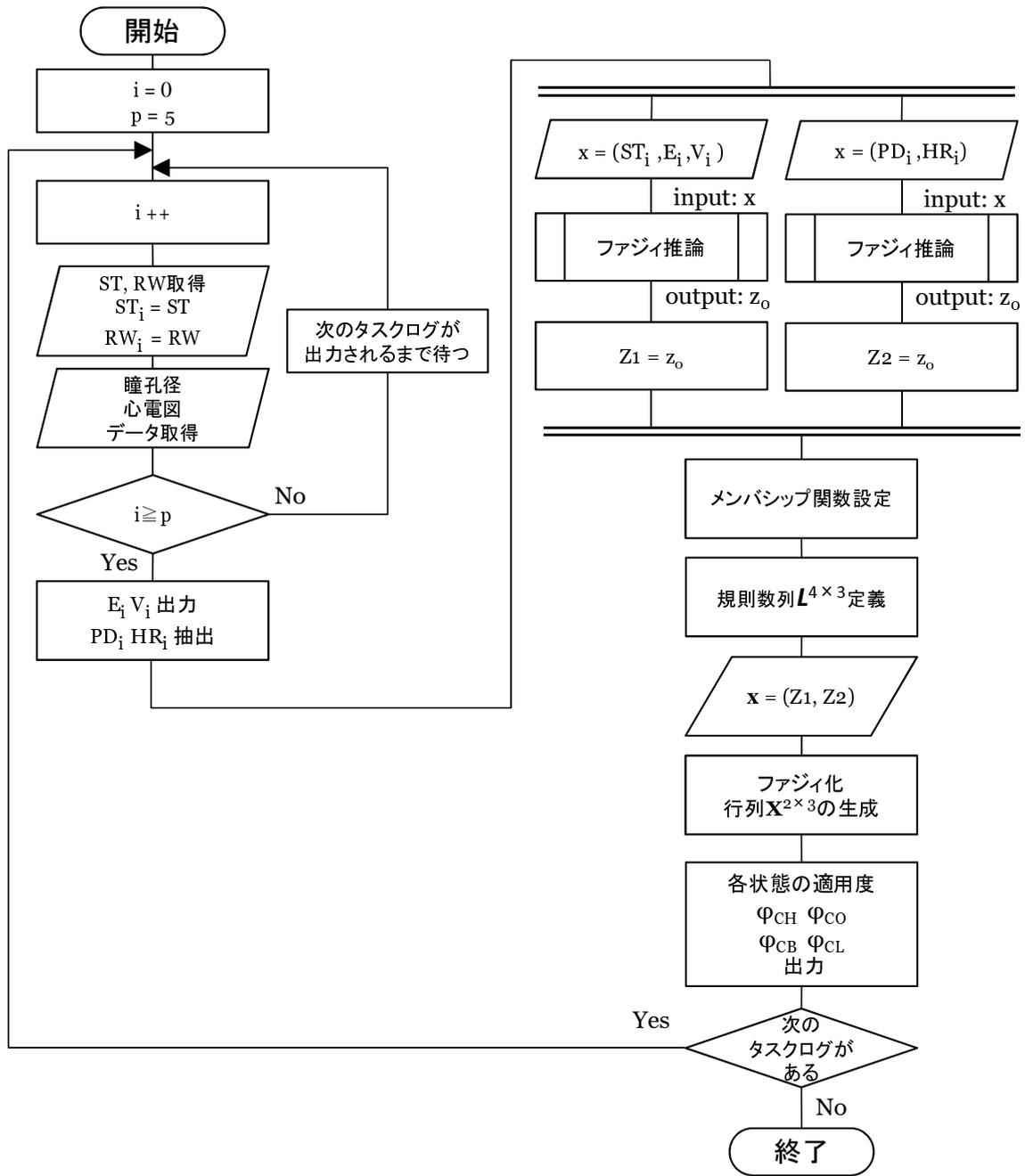


図 3.15: 認知・作業状態推定アルゴリズム

第 4 章 提案手法の妥当性評価実験

4.1 目的

本実験では、まず 3.5 節や 3.6 節で述べた各特徴量が提案手法の入力として妥当であるかを評価し、次に本手法が認知・作業状態を推定する上でパフォーマンス評価および認知負荷評価が妥当であることの確認を目的とした。本実験では前提として、タスク難易度の上昇によって解答時間、分散性、エラー率、瞳孔径及び心拍数が増加し、その結果パフォーマンスが低下、認知負荷が増加すると仮定した。この仮定のもとで、難易度の異なる複数のタスクを実施し、各特徴量や提案手法によって推定されたパフォーマンス及び認知負荷が仮定を満たすかを確認した。

4.2 実験方法

4.2.1 実験環境

実験は、総合研究 11 号館 105 号室にて実施した。実験室のレイアウトを図 4.1 に示す。

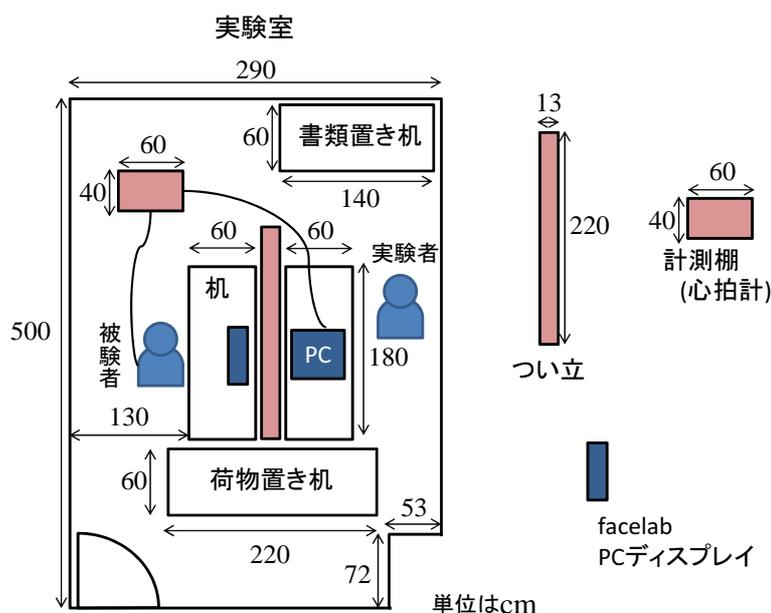


図 4.1: 実験室のレイアウト

計測棚には心拍計や電極装着時の皮質除去のための無水エタノールやガーゼなどを収納しており，脚にキャスターを装着し，被験者の座る椅子の位置に合わせて自由に移動できるようにした．実験中は室温を $24 \pm 1^\circ\text{C}$ ，騒音は 45dB 以下に調節した．また，机上面照度が $525 \pm 50\text{lux}$ ，被験者の眼球付近の照度が $400 \pm 50\text{lux}$ になるように照明及びPCディスプレイ輝度を調節した．

4.2.2 測定手法

心電図計測には株式会社デジテック研究所 Polymate AP216 を使用した．電極には入力インピーダンスが $300\text{G}\Omega$ のアクティブ電極を使用し，測定時には部屋の湿度を $50 \pm 5\%$ に保ちつつ装着部位のアルコール消毒および皮質除去を行い，電極と皮膚間インピーダンスが $5\text{k}\Omega$ 以下になるよう調整した．ハイパスフィルタの時定数を 3.0sec に設定し，ローパスフィルタのカットオフ周波数を 100Hz に設定した．また，商用電源からのハムノイズ除去のため，ノッチフィルタの周波数を 60Hz に設定した．右耳朶及び左耳朶にはそれぞれグランド電極及びリファンレンス電極を装着した．心電図の正極 (ECG+) は右首筋，負極 (ECG-) は左脇腹に装着した．

瞳孔径計測は Seeing Machines, Inc. の赤外線視線計測カメラ faceLAB 5 を用いて行った．計測の際は図 4.2 のように被験者を椅子に座らせ，カメラを机から上方 36° に固定し，被験者の顔を撮影できるように椅子の高さ及び位置を調整した．

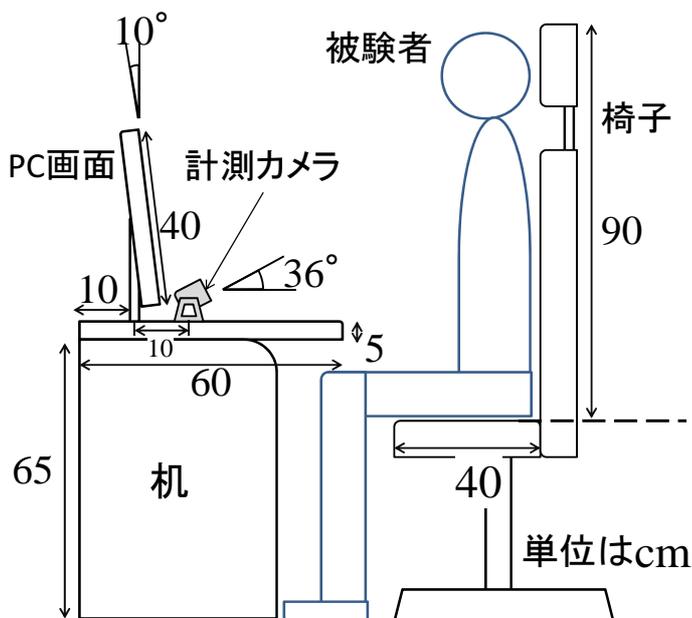


図 4.2: 被験者と計測機材の位置関係

4.2.3 実験スケジュール

実験は平成 25 年 11 月 4 日～21 日の期間で，被験者 1 人あたり 2 日間行った．初日の実験スケジュールは以下の表 4.1 に示す通りである．

表 4.1: 初日の実験スケジュール

所要時間	内容
5 分	瞳孔径測定 (スクリーニング)
10 分	2 桁暗算・4 桁暗算タスク練習
15 分	実験説明・同意確認

合計：30 分

スクリーニングは瞳孔径計測の精度を基準とした．瞳孔径計測カメラは，全ての被験者に対して高い精度で瞳孔径をトラッキングできるとは限らず，コンタクトレンズやメガネを着用している場合，レンズの度数などによっては著しくトラッキング精度が落ちる場合がある．トラッキング精度が低い被験者は正常な計測データを取得できないため，事前にスクリーニングを実施する必要があった．瞳孔径計測の精度確認は，被験者を開眼安静状態で図 4.2 のように座らせ，カメラに瞳孔径をトラッキングさせるだけで良いため，椅子の位置調整も含めて 5 分程度で終了した．このとき，開眼であるのにも関わらずトラッキングエラーが生じる被験者は初日で実験を終了した．トラッキングが正常に行われた被験者には引き続き，実験 2 日目に行う 2 桁暗算及び 4 桁暗算タスクの練習を各 5 分間行わせた．その後実験説明を 15 分程度行い，生理指標計測についてインフォームドコンセント及び同意書を取り，被験者に活動量計測器と表 4.2 に示すような項目を聞く事前アンケート用紙を渡した．

表 4.2: 事前アンケート項目

項目	回答方法
参加日前日の就寝時間と当日の起床時間	自由記述式
朝食の有無	(有・無)からの選択式
昼食の有無	(有・無)からの選択式
食事の内容と時間	自由記述式
現在の体調	快調・普通・不調
体調が不調の場合, その内容	自由記述式
前日の睡眠具合	良・やや良・普通・やや悪・悪
当日の目覚めの良さ	良・やや良・普通・やや悪・悪
服用中の薬	自由記述式
前日から当日までにした運動の内容	自由記述式
カフェイン・アルコールの摂取	(有・無)からの選択式
摂取した場合, 内容と時間	自由記述式
普段の喫煙の量	自由記述式

被験者には実験日2日目の前日に事前アンケート用紙に回答を記入し、同時に活動量計測器を装着して6時間以上睡眠を取り、カフェインやアルコール等の摂取を控え、実験当日まで計測器の装着を維持するよう教示を与えた。このような教示を与えたのは、睡眠時間や生活習慣の統制を意図したからである。生理指標は生活習慣から影響を受けるため、実験当日に寝不足による眠気や疲労の影響を除外する必要がある。したがって、本実験では事前アンケートや活動量計測器によって被験者の生活を統制した。

2日目の実験スケジュールは表4.3の通りである。

表 4.3: 2 日目の実験スケジュール

所要時間	内容
10 分	電極装着・瞳孔径計測カメラ設定
1 分	1 桁暗算加算
1 分	ベースライン測定*
3 分	finger tapping*
1 分	ベースライン測定*
10 分	タスク A**
5 分	NASA-TLX
1 分	ベースライン測定*
10 分	タスク B**
5 分	NASA-TLX
5 分	計器取り外し
合計 : 52 分	

* : 生理指標計測

** : 生理指標計測及びタスクログ取得

内容に記載されている「*」や「**」は、前者が生理指標計測、後者が生理指標計測及びタスクログ取得を行った項目である。表中のタスク A 及びタスク B は 2 桁暗算と 4 桁暗算のいずれかであり、順序効果を打ち消すために被験者によってランダムに割り振り、タスク A が 2 桁暗算でタスク B が 4 桁暗算のグループとその逆の割り振りのグループの被験者数が同程度になるよう調整した。各タスク終了後に行った NASA-TLX はタスクに対するモチベーションや精神負荷などに関する主観指標である [42]。

finger tapping (以降は finger と記す) は暗算加算タスクの対照タスクとして採用した。対照タスクを実施した目的は生理指標計測中、タスク実行時の指や腕の動きなどから生じる認知負荷に影響を持たないアーチファクトが計測され、認知負荷推定の結果に及ぼす影響を確認するためである。したがって、このタスクでの心拍数及び瞳孔径データを 2 桁暗算及び 4 桁暗算と同様、解析の対象として計測した。finger 中は、3.4 項で述べた認知タスクの桁数を 1 桁とした時と同じ画面を表示した。この時、被験者には「最初の呈示数字を見ずに「E」ボタンを押し、計算せずにランダムな 1 桁の数字を解答し、「E」ボタンを押しして解答する」という、知的能力を必要とする認知プロセスを除

去した作業を行うよう教示を与えた。

また、1桁暗算加算は実験開始時にしばしば見られる被験者のモチベーション上昇が評価対象であるタスク A 及びタスク B へ影響することを防ぐためのダミータスクであり、この時のタスクログ及び生理指標を取得しなかった。同様に実験終了間際にしばしば見られるモチベーション上昇による影響を打ち消すため、初日に行った実験説明では2日目の実験スケジュールとして表 4.3 ではなく、この表の「計器取り外し」項目の前に15分のタスク C 実施が追加された表を用いて説明を行った。当日の実験ではタスク B 及び NASA-TLX 終了後に、タスク C はダミータスクであったことを説明し、タスクを実施せず実験を終了した。

4.2.4 被験者

被験者の募集は京都大学生協で行い、(1) 過去に心臓や眼球に疾患の経験がないこと、(2) 母国語が日本語であること、(3) 性別が男性であることを条件とし、先述の実験スケジュールや計測方法を告知した上で、実験参加が可能な被験者を募集した。(1) を提示したのは、過去の疾患によって健常者と異なる瞳孔径や心拍数の変動が計測された場合、提案手法の妥当性評価が正確に行えないため、疾患の経験がある被験者を除外する必要があったためである。(2) については実験説明や被験者の生活統制の教示を正確に行うためである。最後の(3)に関しては、被験者への電極装着時に不快感を与えないように、実験者と性別を一致させる必要があったためである。

募集の結果、31名の男性大学生が本実験に参加した。被験者の属性を表 4.4 に示す。

表 4.4: 被験者属性

ID	年齢	メガネ着用	初日	時間帯	2 日目	時間帯	タスク A
s01	21	無	2013/11/4	10:00-10:30	2013/11/12	8:45-10:15	2 桁暗算
s02	22	無	2013/11/4	10:30-11:00	2013/11/14	13:00-14:30	2 桁暗算
s03	20	無	2013/11/4	11:00-11:30	2013/11/14	8:45-10:15	2 桁暗算
s04	25	有	2013/11/4	13:00-13:30	2013/11/12	14:45-16:15	4 桁暗算
s05	18	無	2013/11/4	14:00-14:30	2013/11/12	16:30-18:00	2 桁暗算
s06	18	無	2013/11/4	15:00-15:30	2013/11/12	10:30-12:00	4 桁暗算
s07	19	無	2013/11/4	16:00-16:30	2013/11/16	8:45-10:15	4 桁暗算
s08	20	有	2013/11/5	9:00-9:30	2013/11/14	10:30-12:00	4 桁暗算
s09	22	有	2013/11/5	10:00-10:30	2013/11/19	10:30-12:00	2 桁暗算
s10	20	無	2013/11/5	11:00-11:30	2013/11/12	13:00-14:30	2 桁暗算
s11	20	有	2013/11/5	15:30-16:00	2013/11/13	13:00-14:30	2 桁暗算
s12	23	無	2013/11/6	13:30-14:00	2013/11/13	14:45-16:15	4 桁暗算
s13	23	無	2013/11/6	14:30-15:00	2013/11/11	14:45-16:15	2 桁暗算
s14	19	無	2013/11/6	15:00-15:30	2013/11/12	18:15-19:45	4 桁暗算
s15	24	無	2013/11/7	9:00-9:30	2013/11/13	10:30-12:00	4 桁暗算
s16	21	無	2013/11/7	11:30-12:00	2013/11/13	8:45-10:15	2 桁暗算
s17	25	無	2013/11/7	13:00-13:30	2013/11/11	13:00-14:30	4 桁暗算
s18	20	無	2013/11/7	13:30-14:00	2013/11/14	16:30-18:00	4 桁暗算
s19	21	無	2013/11/7	14:00-14:30	2013/11/11	8:45-9:45	2 桁暗算
s20	19	無	2013/11/7	14:30-15:00	2013/11/15	10:30-12:00	4 桁暗算
s21	23	無	2013/11/7	16:00-16:30	2013/11/21	14:45-16:15	4 桁暗算
s22	20	無	2013/11/8	10:00-10:30	2013/11/18	14:45-16:15	4 桁暗算
s23	22	無	2013/11/8	11:00-11:30	2013/11/11	16:30-18:00	4 桁暗算
s24	22	無	2013/11/8	11:30-12:00	2013/11/19	13:00-14:30	4 桁暗算
s25	22	無	2013/11/8	13:00-13:30	2013/11/15	13:00-14:30	2 桁暗算
s26	20	無	2013/11/8	17:00-17:30	2013/11/15	16:30-18:00	2 桁暗算
s27	23	無	2013/11/11	10:30-11:00	2013/11/15	14:45-16:15	4 桁暗算
s28	20	無	2013/11/12	12:15-12:45	2013/11/15	8:45-10:15	2 桁暗算
s29	23	無	2013/11/14	15:15-15:45	2013/11/20	10:30-12:00	2 桁暗算
s30	20	無	2013/11/14	15:45-16:15	2013/11/18	10:30-12:00	2 桁暗算
s31	22	有	2013/11/6	15:30-16:00	-	-	-

また、タスク A の項目には 2 日目の実験日にてタスク A で行ったタスク名を記した。被験者 s31 は初日のスクリーニングで著しい瞳孔径トラッキングエラーが認められたため、初日の参加のみであった。トラッキングエラーの原因として、被験者 s31 の着用していたメガネのレンズが厚く、瞳孔径が識別困難であったことが挙げられる。他にも数名の被験者がメガネを着用していたが、トラッキングは比較的安定していたため 2 日目も実験に参加してもらった。

4.3 実験結果

2 日目に参加した被験者は表 4.4 に示した s1~s30 の 30 名であった。この 30 名のうち、1 名 (被験者 ID : s06) の ECG 波形において、図 4.3 に表すような R 波減高かつ RR 間隔が明らかに短い異常波形が見られた。

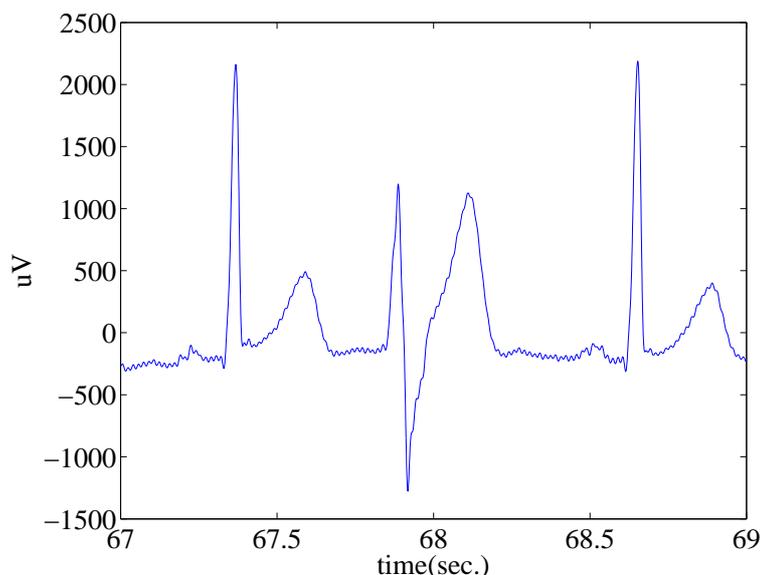


図 4.3: 被験者 s06 の心電図データ

ノイズや体動によるアーチファクトは認められず、正常な ECG 波形も多く計測されているため、異常波形は計測上のミスが原因であるとは考え難く、当該被験者の心拍の特性によるものであると考えられたため、本研究では被験者 s06 を解析対象外とした。また、被験者 s16 及び s24 は実験中に体調不良及び強い眠気を訴えたため、この 2 名のデータも解析対象外とした。最終的に解析対象となった被験者は以上の被験者を除いた 27 名となった。

4.3.1 NASA-TLX

全被験者における，2桁暗算及び4桁暗算でのNASA-TLXによるメンタルワークロードの平均を表4.5及び図4.4に示す。

表 4.5: NASA-TLX によるメンタルワークロード

2桁暗算	4桁暗算
55.2 (9.3)	64.1 (7.0)

() 内は標準偏差

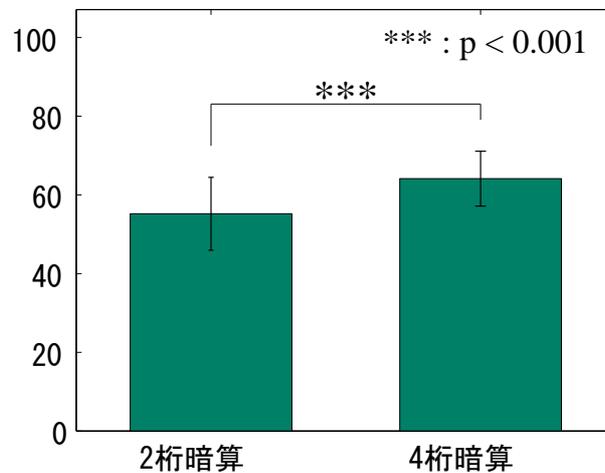


図 4.4: メンタルワークロードのグラフ及び検定結果

対のあるt検定を行ったところ，有意差が認められたため ($p \leq 0.001$)，本研究の認知タスクの難易度設定は被験者に異なる認知負荷を与える点で，有効であったと言える。

4.3.2 解答時間，分散性及びエラー率

全被験者の各タスクについて3.5.1項で抽出した各特徴量の平均値と標準偏差をそれぞれ表4.6に示し，被験者ごとの特徴量(解答時間，分散性及びエラー率)の平均値を表4.7に示す。

さらに，対のあるt検定をかけた結果のグラフを図4.5，図4.6及び図4.7に示す。

表 4.6: 各タスクの平均特徴量 (解答時間・分散性・エラー率)

解答時間		分散性		エラー率	
2桁暗算	4桁暗算	2桁暗算	4桁暗算	2桁暗算	4桁暗算
4.13 (1.33)	11.24 (2.80)	1.03 (1.06)	10.72 (8.23)	3.81 (3.08)	16.27 (12.65)

() 内は標準偏差

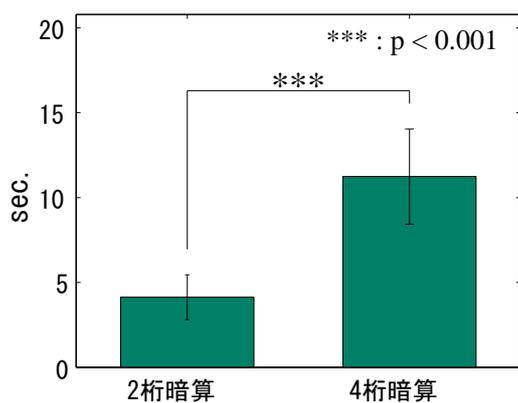


図 4.5: 各タスクごとの平均解答時間

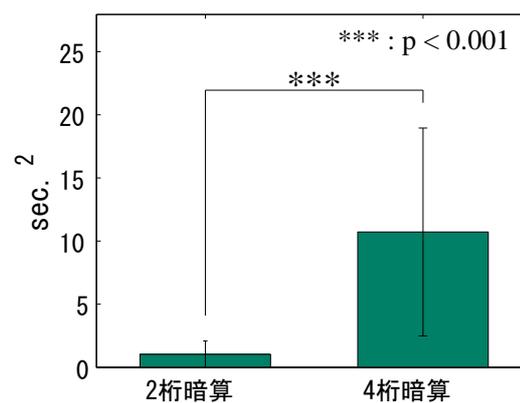


図 4.6: 各タスクごとの平均分散性

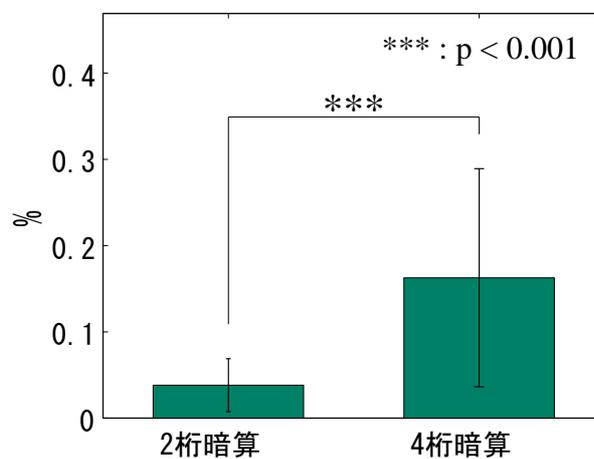


図 4.7: 各タスクごとの平均エラー率

表 4.7: 各被験者の平均特徴量 (解答時間, 分散性, エラー率)

被験者	解答時間 (sec.)		分散性 (sec. ²)		エラー率 (%)	
	2桁暗算	4桁暗算	2桁暗算	4桁暗算	2桁暗算	4桁暗算
s01	3.28	11.96	0.43	18.41	7.69	35.50
s02	3.7	9.02	0.26	3.08	2.81	7.41
s03	4.07	13.56	0.96	16.48	2.50	30.00
s04	5.03	16.91	1.44	21.13	2.13	5.93
s05	4.14	12.02	1.37	14.90	9.82	23.68
s07	4.12	9.34	0.55	3.84	0.00	12.16
s08	4.2	11.23	1.30	31.29	4.50	10.95
s09	3.64	13.8	0.98	28.02	2.29	19.41
s10	3.46	8.62	0.51	3.85	5.37	11.48
s11	3.46	10.65	0.33	9.43	5.00	18.18
s12	3.89	10.16	0.90	5.04	3.36	14.47
s13	3.37	8.34	0.64	3.90	0.00	1.43
s14	3.69	11.88	0.29	8.39	1.55	6.67
s15	4.12	9.2	1.11	2.99	7.83	12.55
s17	2.84	10.87	0.42	8.70	9.82	61.90
s18	4.35	13.41	1.14	6.84	0.00	2.78
s19	3.35	10.56	0.61	4.78	9.50	22.22
s20	8.36	17.56	4.99	16.25	3.16	15.56
s21	7.93	16.29	2.64	24.23	3.61	18.46
s22	3.8	11.78	0.59	7.97	0.00	4.88
s23	2.69	5.99	0.13	1.66	0.57	6.05
s25	3.8	8.95	0.65	7.36	4.76	20.38
s26	3.82	11.94	0.55	5.78	0.00	3.90
s27	3.04	6.35	0.24	1.85	2.72	7.69
s28	3.35	10.47	0.44	16.86	6.43	25.22
s29	4.03	11.22	1.05	7.39	4.35	25.50
s30	5.9	11.48	3.32	9.02	3.21	15.00

各特微量に対して対のある t 検定を行ったところ、有意差が認められた ($p < 0.001$).

4.3.3 心拍数及び瞳孔径

各タスクについて 3.6.1 項で抽出した各特微量 (瞳孔径及び心拍数) の平均値と標準偏差をそれぞれ表 4.8 に示す.

表 4.8: 各タスクの平均特微量 (瞳孔径・心拍数)

瞳孔径 (mm)			心拍数 (bpm)		
finger	2桁暗算	4桁暗算	finger	2桁暗算	4桁暗算
3.87 (0.44)	4.01 (0.47)	4.24 (0.55)	73.8 (13.0)	75.09 (11.7)	77.18 (12.5)

() 内は標準偏差

各特微量について、finger, 2桁暗算及び4桁暗算を水準として分散分析を行った結果、瞳孔径では $F(2,52)=24.90$, $MSe=4.45 \times 10^{-2}$, $p < 0.001$, 心拍数では $F(2,52)=8.85$, $MSe=9.18$, $p < 0.001$ となり、各特微量ともにタスクによる効果が認められた. さらに各特微量に対して tukey 多重検定を行った. 結果のグラフを図 4.8 及び図 4.9 に示す.

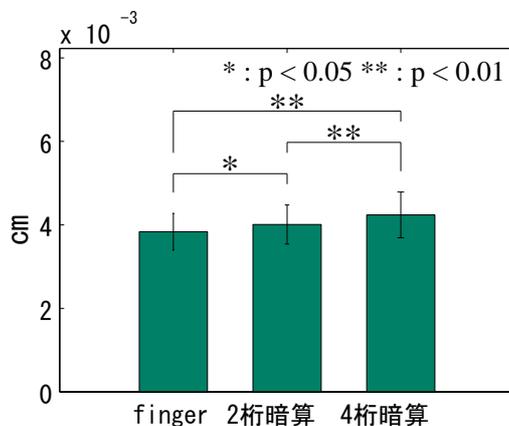


図 4.8: 各タスクごとの平均瞳孔径

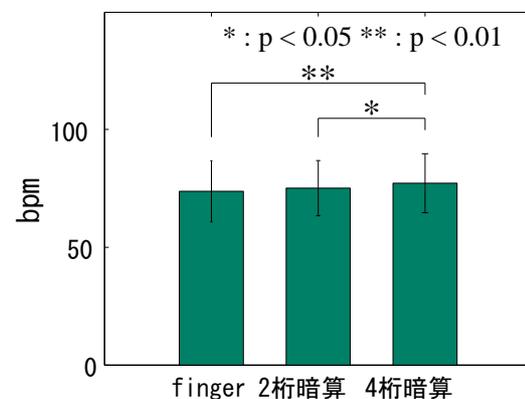


図 4.9: 各タスクごとの平均心拍数

瞳孔径に関しては各水準で有意差が見られたが、心拍数では finger と 2桁暗算で有意差が見られなかった.

また、最後に各被験者ごとの特微量の平均値を表 4.9 に示す.

表 4.9: 各被験者の平均特徴量 (瞳孔径, 心拍数)

被験者	瞳孔径 (mm)			心拍数 (bpm)		
	finger	2 桁暗算	4 桁暗算	finger	2 桁暗算	4 桁暗算
s01	3.45	3.62	3.89	84.50	86.64	90.79
s02	3.61	3.87	4.64	68.67	72.50	77.00
s03	3.34	3.46	3.75	83.33	88.25	85.42
s04	4.19	3.50	4.00	78.04	74.89	78.73
s05	3.82	4.39	4.35	75.96	74.29	73.91
s07	3.61	3.78	3.91	61.61	63.13	66.32
s08	4.91	4.38	4.76	64.38	70.13	70.32
s09	3.84	4.06	4.24	73.75	72.25	74.93
s10	3.56	3.67	3.95	61.49	65.56	72.57
s11	3.79	4.06	4.36	83.49	81.36	81.48
s12	3.61	3.56	3.65	99.35	101.72	104.57
s13	3.53	3.84	3.97	65.64	65.70	66.02
s14	3.97	4.22	4.34	51.61	60.45	59.00
s15	3.55	3.79	3.72	57.93	58.10	58.89
s17	3.43	3.60	3.78	64.56	63.18	62.59
s18	4.18	5.10	5.57	87.64	78.21	85.95
s19	2.89	3.25	3.28	80.44	87.31	87.87
s20	4.25	4.75	4.79	83.52	88.99	98.63
s21	3.75	3.71	3.83	78.24	83.19	84.22
s22	3.88	4.40	4.60	100.94	91.77	99.13
s23	3.32	3.44	3.52	65.75	71.18	68.60
s25	3.96	4.39	4.67	79.05	83.16	84.91
s26	4.04	3.98	4.05	59.60	61.04	64.52
s27	3.83	3.79	3.95	65.53	71.31	70.02
s28	4.64	4.71	5.07	92.62	88.20	85.21
s29	4.11	4.28	5.08	61.31	60.99	68.16
s30	4.54	4.62	4.76	62.07	64.04	64.24

4.3.4 パフォーマンス、認知負荷、認知・作業状態の推定結果

次に、パフォーマンス推定及び認知負荷推定について、タスクごとの平均値及び標準偏差を表 4.10 に示す。パフォーマンス推定について、対のある t 検定を行ったところ、有意差が認められた ($p < 0.001$)。グラフ及び検定結果を図 4.10 に示す。また、認知負荷推定において finger, 2桁暗算及び4桁暗算を水準として分散分析を行った結果、 $F(2,52)=26.42$, $MSe=0.01$, $p < 0.001$ となった。さらに、tukey 多重検定を行った結果を図 4.11 に示す。

表 4.10: 各タスクの平均パフォーマンス及び平均認知負荷

パフォーマンス		認知負荷		
2桁暗算	4桁暗算	finger	2桁暗算	4桁暗算
0.71 (0.21)	0.33 (0.22)	0.38 (0.14)	0.46 (0.13)	0.58 (0.13)

() 内は標準偏差

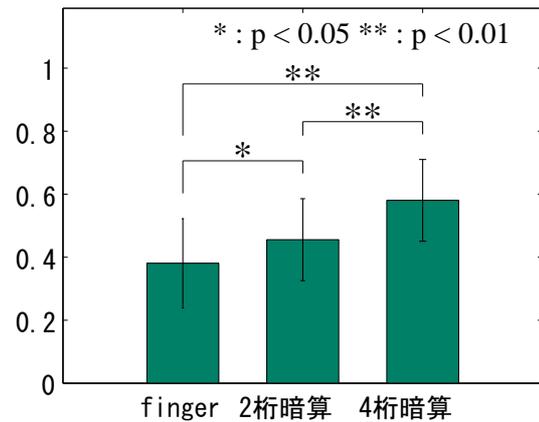
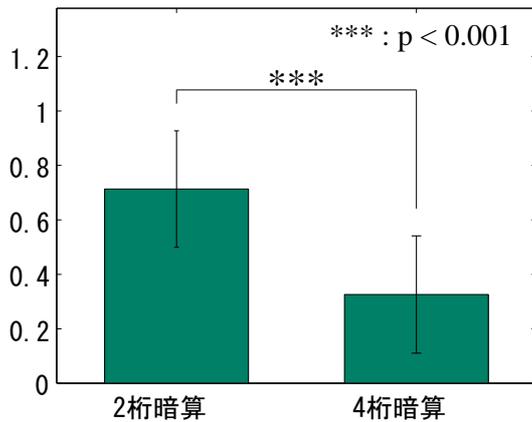


図 4.10: 各タスクごとの平均パフォーマンス 図 4.11: 各タスクごとの平均認知負荷

また、各推定の一例として被験者 s10 の 4桁暗算時の推定結果を各入力特徴量と共に図 4.12 及び図 4.13 に示す。グラフ上に引かれた赤線は 3.5.2 項や 3.6.2 項で設定した閾値 T_{low} , T_{Middle} 及び T_{High} である。

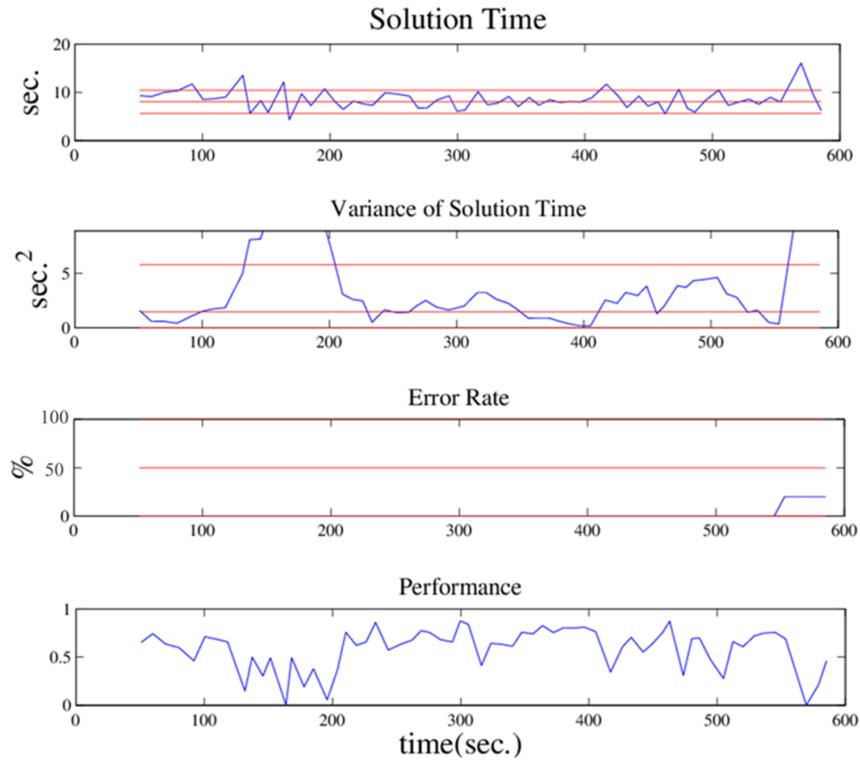


図 4.12: 解答時間, 分散性, エラー率及びパフォーマンス推定結果

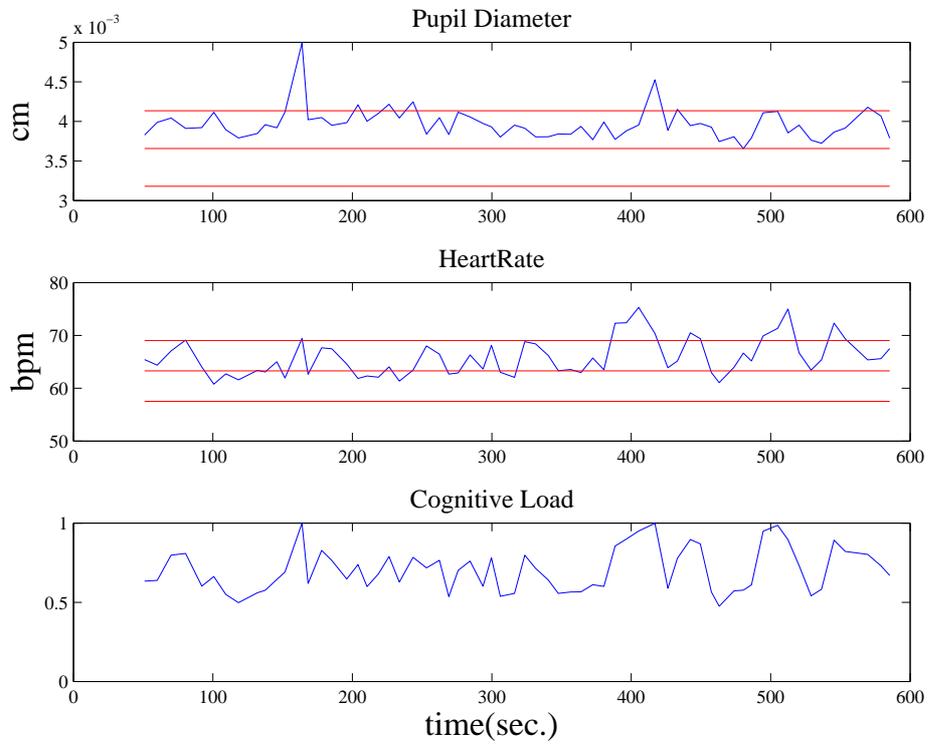


図 4.13: 瞳孔径, 心拍数及び認知負荷推定結果

また，各被験者ごとのパフォーマンス及び認知負荷の平均値を表 4.11 に示す．

表 4.11: 各被験者の平均パフォーマンス及び平均認知負荷

被験者	パフォーマンス		認知負荷		
	2桁暗算	4桁暗算	finger	2桁暗算	4桁暗算
s01	0.84	0.16	0.39	0.36	0.60
s02	0.89	0.56	0.37	0.48	0.76
s03	0.69	0.14	0.32	0.51	0.58
s04	0.56	0.04	0.53	0.31	0.57
s05	0.61	0.18	0.69	0.76	0.85
s07	0.77	0.51	0.11	0.33	0.53
s08	0.63	0.26	0.65	0.65	0.66
s09	0.75	0.17	0.36	0.62	0.51
s10	0.82	0.54	0.39	0.54	0.67
s11	0.87	0.26	0.53	0.41	0.63
s12	0.75	0.36	0.40	0.46	0.40
s13	0.84	0.58	0.38	0.51	0.71
s14	0.88	0.15	0.35	0.38	0.48
s15	0.68	0.52	0.26	0.38	0.34
s17	0.84	0.17	0.45	0.24	0.55
s18	0.62	0.26	0.45	0.36	0.37
s19	0.79	0.37	0.38	0.33	0.41
s20	0.11	0.04	0.25	0.43	0.73
s21	0.13	0.05	0.37	0.64	0.76
s22	0.81	0.26	0.22	0.27	0.49
s23	0.97	0.83	0.14	0.53	0.53
s25	0.81	0.48	0.37	0.32	0.64
s26	0.81	0.28	0.63	0.58	0.54
s27	0.92	0.82	0.23	0.50	0.55
s28	0.84	0.34	0.23	0.34	0.49
s29	0.66	0.27	0.41	0.57	0.78
s30	0.37	0.20	0.43	0.50	0.54

次に、認知・作業状態について推定結果を評価した。各タスクの認知・作業状態の平均適応度を表 4.12 に示す。

表 4.12: 各タスクの認知・作業状態の平均適応度

High W.S.		Optimal W.S.		Bad W.S.		Low W.S.	
2桁暗算	4桁暗算	2桁暗算	4桁暗算	2桁暗算	4桁暗算	2桁暗算	4桁暗算
0.24 (0.12)	0.16 (0.13)	0.37 (0.18)	0.13 (0.09)	0.12 (0.11)	0.36 (0.18)	0.11 (0.10)	0.22 (0.11)

() 内は標準偏差

また、それぞれの状態の適応度について対のある t 検定を行った。その結果を図 4.14 ~ 図 4.17 に示す。

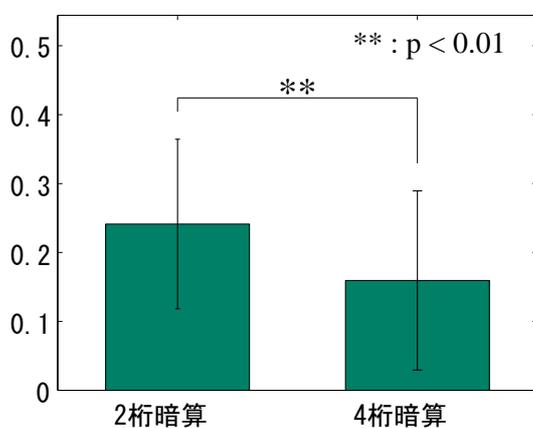


図 4.14: High W.S. の平均適応度

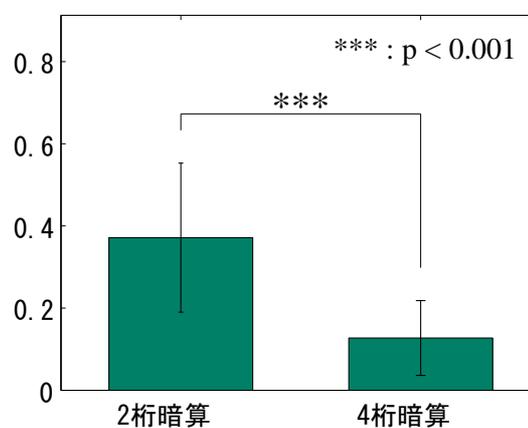


図 4.15: Optimal W.S. の平均適応度

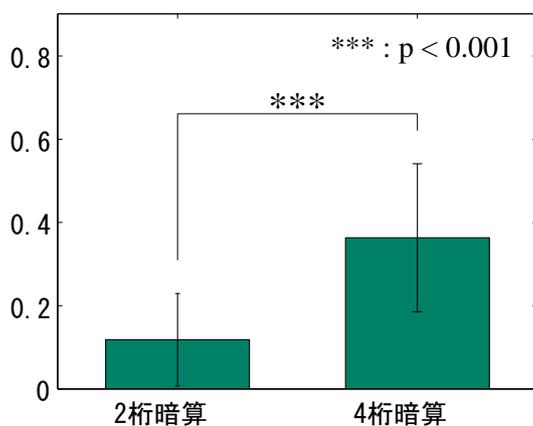


図 4.16: Bad W.S. の平均適応度

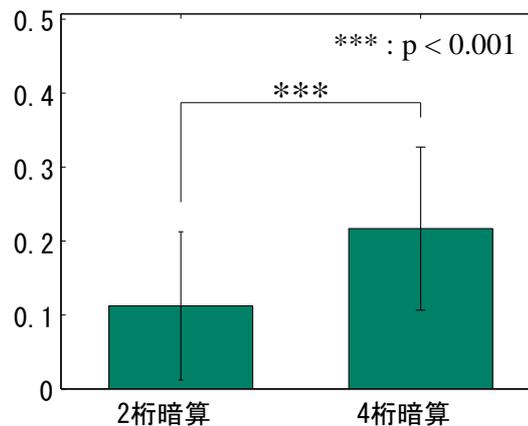


図 4.17: Low W.S. の平均適応度

2桁暗算では、4桁暗算より High W.S. 及び Optimal W.S. の適応度が高い、つまりそれぞれの状態が被験者に現れやすく、逆に Bad W.S. 及び Low W.S. が現れにくいという結果になった。

最後に、被験者ごとの認知・作業状態の平均適応度を表 4.13 に示す。

表 4.13: 各被験者の認知・作業状態の平均適応度

被験者	High W.S.		Optimal W.S.		Bad W.S.		Low W.S.	
	2桁暗算	4桁暗算	2桁暗算	4桁暗算	2桁暗算	4桁暗算	2桁暗算	4桁暗算
s01	0.17	0.08	0.55	0.05	0.06	0.43	0.06	0.23
s02	0.32	0.44	0.38	0.09	0.04	0.29	0.04	0.07
s03	0.28	0.06	0.32	0.06	0.14	0.47	0.11	0.26
s04	0.13	0.02	0.38	0.02	0.09	0.46	0.27	0.31
s05	0.45	0.09	0.09	0.04	0.26	0.75	0.06	0.06
s07	0.16	0.23	0.57	0.24	0.07	0.26	0.09	0.19
s08	0.39	0.13	0.17	0.10	0.21	0.53	0.10	0.17
s09	0.42	0.07	0.21	0.08	0.14	0.36	0.06	0.34
s10	0.39	0.36	0.30	0.14	0.07	0.28	0.06	0.10
s11	0.26	0.12	0.49	0.11	0.05	0.50	0.04	0.18
s12	0.27	0.12	0.38	0.21	0.11	0.20	0.09	0.38
s13	0.34	0.41	0.34	0.12	0.07	0.25	0.06	0.09
s14	0.19	0.06	0.53	0.07	0.04	0.32	0.04	0.36
s15	0.20	0.13	0.42	0.38	0.10	0.15	0.17	0.27
s17	0.09	0.07	0.70	0.06	0.04	0.40	0.06	0.28
s18	0.15	0.08	0.42	0.13	0.12	0.19	0.20	0.44
s19	0.16	0.11	0.56	0.23	0.07	0.24	0.09	0.31
s20	0.04	0.02	0.05	0.02	0.27	0.68	0.47	0.14
s21	0.05	0.03	0.04	0.03	0.53	0.72	0.21	0.12
s22	0.11	0.11	0.63	0.12	0.05	0.30	0.09	0.33
s23	0.39	0.36	0.31	0.30	0.02	0.07	0.02	0.05
s25	0.13	0.31	0.59	0.14	0.06	0.31	0.08	0.14
s26	0.42	0.13	0.24	0.12	0.08	0.36	0.06	0.26
s27	0.35	0.41	0.35	0.27	0.03	0.07	0.03	0.06
s28	0.16	0.14	0.57	0.17	0.05	0.27	0.07	0.28
s29	0.32	0.13	0.26	0.05	0.17	0.55	0.11	0.15
s30	0.17	0.09	0.17	0.08	0.28	0.38	0.29	0.27

また、各推定の例として被験者 s10 の 2桁暗算及び 4桁暗算時の推定結果をそれぞれ

図 4.18 及び図 4.19 に示し、被験者 s22 の 2 桁暗算及び 4 桁暗算時の推定結果を図 4.20 及び図 4.21 に示す。

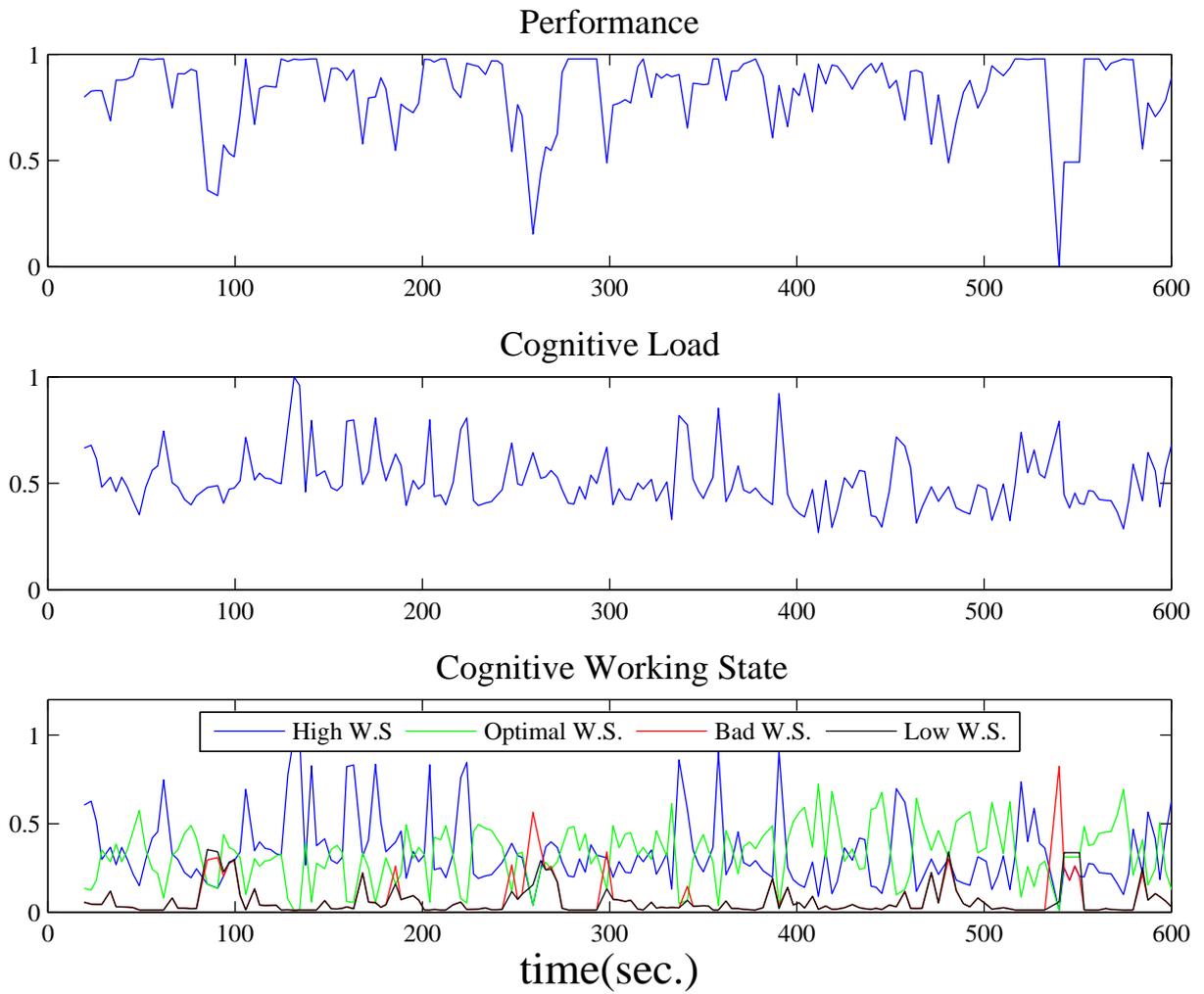


図 4.18: 被験者 s10 の認知・作業状態推定結果 (2 桁暗算)

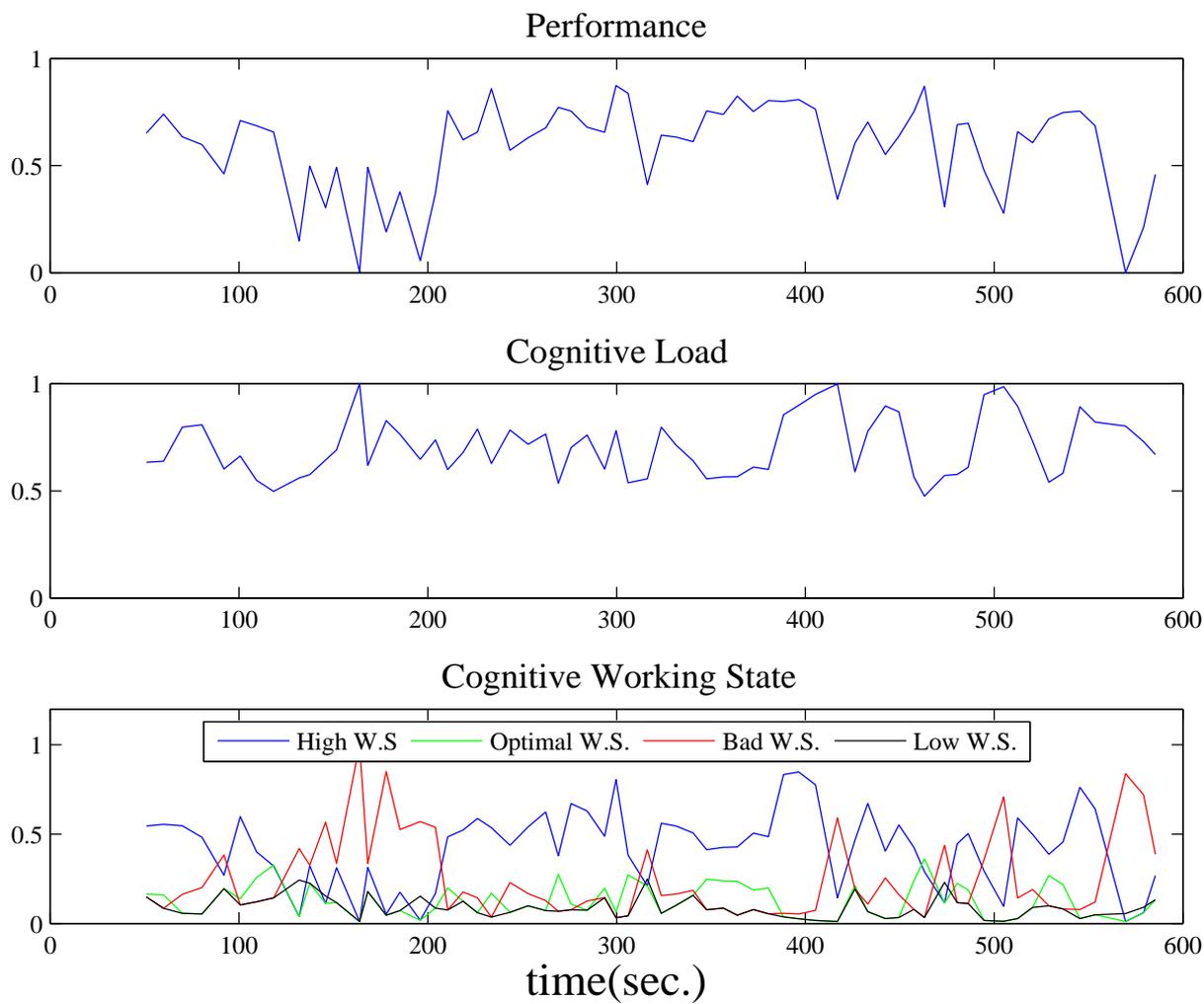


図 4.19: 被験者 s10 の認知・作業状態推定結果 (4 桁暗算)



図 4.20: 被験者 s22 の認知・作業状態推定結果 (2 桁暗算)

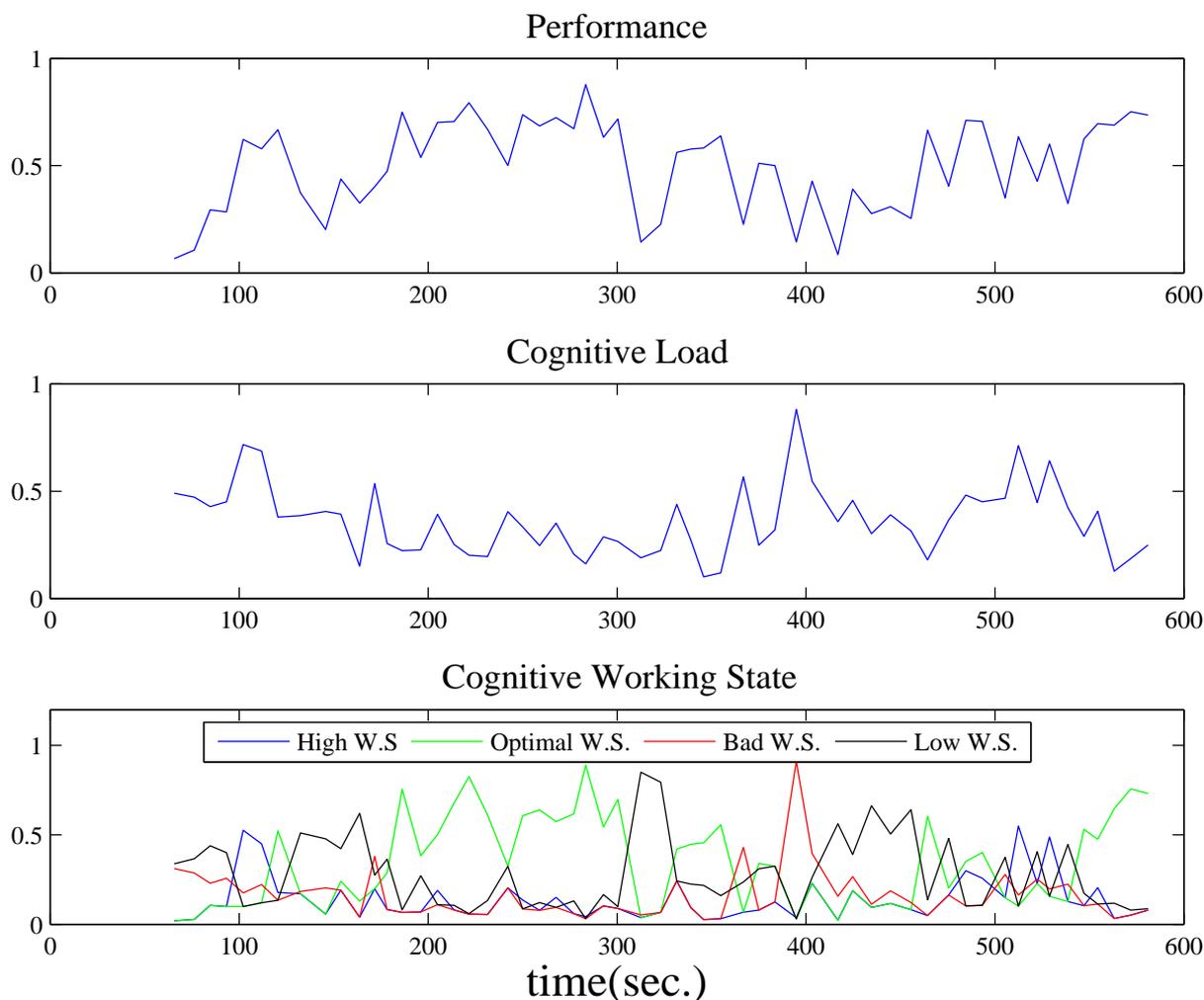


図 4.21: 被験者 s22 の認知・作業状態推定結果 (4桁暗算)

4.4 考察

4.4.1 特徴量の検討

本実験での想定として暗算加算タスクでは、桁数の増加に伴って短期記憶にかかる負荷も増加するため、記憶の保持や想起も困難となる。これにより表 3.2 で記したようなタスクの認知ステップの他にも、追加で新たな認知ステップ (記憶情報の更新や思い出しなど) が生じ、解答時間や分散性が増加する可能性がある。また、想起失敗によるエラー率の増加も起こり得る。同様に、要求される認知負荷も増加するため、瞳孔径及び心拍数が増加すると考えられる。4.3.2 項及び 4.3.3 項で述べた結果はこれらの仮定を満たしているため、各特徴量がタスク難易度によるパフォーマンス及び認知負荷

の変動を表し得る指標であると言える。しかし、図 4.9 で示した tukey 多重検定結果によると、心拍数は finger と 2 桁暗算の間に有意差は認められなかった。短期記憶に負荷をかけるタスクを用いた実験^[22,23,26]ではコントロール群とタスク群で心拍数の有意差が認められているが、用いられたタスクは5チャンク～7チャンクの短期記憶容量を要求するものであり、本研究では4桁暗算と同程度である。tukey 多重検定によると、finger - 4 桁暗算間 ($p < 0.01$) と、2 桁暗算-4 桁暗算間 ($p < 0.05$) では有意差が認められており、既往研究の結果と一致しているため、高認知負荷時に心拍数が有意に上昇することは示唆できるが、2 桁暗算のように比較的認知負荷が低いタスクは上昇値が小さく、その結果 finger - 2 桁暗算間で有意差が現れなかったと考えられる。以上から、心拍数は認知負荷を反映する指標ではあるが変動量が瞳孔径に比べ低く、短期記憶容量5チャンク相当以上の認知負荷下で有意な心拍数の上昇が見られると考えられる。

4.4.2 パフォーマンス推定及び認知負荷推定の検討

パフォーマンス推定では図 4.10 からタスク難易度によるパフォーマンス変動が認められ ($p < 0.001$)、認知負荷推定では分散分析によりタスクの効果が認められ ($F(2,52)=26.42$, $MSe=0.01$, $p < 0.001$)、図 4.11 から finger - 2 桁暗算間では $p < 0.05$ 、finger - 4 桁暗算間及び2桁暗算-4桁暗算間では $p < 0.01$ で有意差が認められた。これらの結果は4.1節で述べた、タスク難易度によるパフォーマンス低下及び認知負荷増加の想定通りであり、提案手法による推定結果が妥当であると考えられる。

しかし、パフォーマンス推定について、各入力特徴量(解答時間など)や推定結果はそれぞれタスク難易度で有意差があると分かったが、本研究のように複数の特徴量を用いてパフォーマンスを評価する必要はなく、いずれか1つのみを用いた評価で十分である可能性は否定できない。これまで、パフォーマンス評価に解答時間やエラー率のみを対象にする知的生産性研究^[2-5]が行われていることから、確かに特徴量が1種類であってもパフォーマンス評価は可能であるように思える。しかし、実際の知的作業の現場では正しく評価されない恐れがある。これまでの研究や本研究における被験者実験では、あくまでも実験参加という形で被験者はタスクを遂行しているため、ホーソン効果^[43]等から一定の高モチベーション下にあり、その点で解答時間などの統制が働いている。実験結果の条件間比較を行う際のモチベーション統制が目的であるなら、ホーソン効果は問題なくむしろ望ましい効果であるが、仮にこの実験条件下で解答時間などの1つの特徴量がパフォーマンス評価に有効であると示唆されても、実際の知的作業の現場においても同様に有効な指標であるとは言い難い。実際の知的作業環境

下では、モチベーションなどの執務者の内的要因は統制されておらず、常に変動している。そのため、執務者が解答時間を意識するあまり正確性を欠くことも考えられる。このとき、従来の評価手法によると解答時間は短いためパフォーマンスが高いと評価されるが、エラー率の高さから真にパフォーマンスが高いとは解釈できない。また、逆にエラー率のみを評価対象とした場合、正確さを意識するあまり解答時間が長引いた時でも、エラー率が低ければパフォーマンスは高いと評価されるが、解答時間の点からパフォーマンス評価が正しいとは言い難い。一方で、解答時間の平均が同程度でも、解答時間の分散が大きい場合と小さい場合ではパフォーマンスに対する解釈が異なると思われる。以上のような問題点から、本研究では3.5節で述べたようにパフォーマンスを「解答時間」、「分散性」及び「エラー率」の3要素から成り立つ知的作業の「質」と定義し、その推定手法を提案した。本実験結果から、提案手法において推定されたパフォーマンスも各特微量と同様にタスク難易度による変動を十分表していることが示唆されたため、パフォーマンス推定に用いる各特微量と推定手法は妥当であると言える。

また同様の議論が、認知負荷推定についても考えられる。瞳孔径は認知負荷を反映するだけでなく、注意や感情、好み、眠気、覚醒度も反映し、さらに対光反射や筋活動によっても変動する^[37,39]。瞳孔径の変動のみからこれらを判別することは困難であり、これまでに明確な判別手法の確立はされていない。一般に瞳孔径を扱った人間の心的活動(注意など)の評価では、光環境を統制したり体動を抑制したりして実験を行う場合が多く、瞳孔径のみを用いた評価手法には限界がある。心拍数に関しても、認知負荷以外に深呼吸や交感・副交感神経の優位関係などによって変動するため、瞳孔径と同様の問題点がある。しかし、これまでの研究^[17-27]でこれら2つの指標が認知負荷と正の相関を持つことが報告されているため、認知負荷変動に対する各指標の変動はほぼ共通していると考えられる。したがって、2指標を同時に用いて認知負荷を推定することで、1指標のみに依存した評価手法に比べてより正確な認知負荷評価が可能になるとと思われる。また、パフォーマンス推定の時と同様に、本実験結果から提案手法で推定した認知負荷も各指標と同様に異なる難易度のタスクによる効果が認められたため、認知負荷に用いる各特微量と推定手法は妥当であったと思われる。

4.4.3 認知・作業状態推定の検討

最後に、認知・作業状態の推定結果について図 4.14 ~ 図 4.17 から、難易度上昇に伴って High W.S. 及び Optimal W.S. は適応度が低下し、Bad W.S. 及び Low W.S. の

適応度は上昇することが分かった。特に、Optimal W.S. では2桁暗算及び4桁暗算でそれぞれ平均値が0.37及び0.13、Bad W.S. ではそれぞれ0.12及び0.36と他の状態に比べて差が比較的大きい。したがって、低難易度であれば、高パフォーマンス低認知負荷、高難易度であれば逆に低パフォーマンス高認知負荷であるという想定が満たされていると言える。

次に、実際の推定結果の一例である図4.18～図4.21において、これらの結果から考えられる被験者の作業状況や認知・作業状態についての考察を述べる。

まず、図4.18に示した被験者s10の2桁暗算時の推定結果について、認知・作業状態はHigh W.S. 及びOptimal W.S. がほぼ全ての作業時間で優位に現れ、特に100秒～250秒の区間や350秒～400秒の区間でHigh W.S. の高いピークが複数形成されており、450秒付近や520秒付近にも高いピークが認められる。このような2状態の優位がよく見られたのは、2桁暗算であるため難易度が低く、想起失敗などのエラーが生じにくいため、作業に集中できていたためであると考えられる。しかし、Optimal W.S. に比べてHigh W.S. の方が適応度が高く現れている傾向にあるため、疲労の蓄積が懸念される。一方で、Low W.S. とBad W.S. に注目するとそれぞれ、100秒、180秒、250秒、480秒、550秒、600秒付近で、他の時間と比べ、やや高い適応度が認められる。これらのピークは、High W.S. が優位である区間の後によく見られ、被験者s10は疲労の蓄積などが原因で作業のペースが低下していると考えられる。しかし、それぞれ2状態のピーク形成後、High W.S. やOptimal W.S. が優位となっているため、作業へ再び集中できていると考えられる。ここで、例えば被験者s10がさらに作業へ集中しやすいように適切な休息のタイミングを提案するのなら、Low W.S. やBad W.S. のピークが認められる区間はもちろんのこと、High W.S. が優位である時間が長い区間(例えば100秒～250秒など)の後に休息を促すと良いと考えられる。

次に、同被験者の4桁加算時の推定結果、図4.19について、先述の2桁暗算の場合とは異なり、High W.S. とBad W.S. が優位である。また、前半100秒～200秒及び後半400秒～600秒ではBad W.S. が現れやすくなっているが、興味深いことに、これらの時間帯での状態遷移を見ると、前半と後半でBad W.S. の生じた原因が異なっている可能性が考えられる。前半では常に4状態の中でBad W.S. が最も高い適応度を示しており、後にHigh W.S. が優位となることから、タスクに慣れるまでに作業への集中が十分でなく、さらに4桁暗算という高難易度のタスクを解答しているため、Bad W.S. が現れやすくなったと考えられる。一方で、後半のBad W.S. のピークについて、400秒以前ではHigh W.S. が優位であることから、High W.S. での疲労蓄積によってBad

W.S. が現れやすくなったと考えられる。さらに、その後 High W.S. と交互にピークの形成が見られることから、疲労の回復が期待できず、Bad W.S. への遷移と High W.S. への復帰を繰り返し、作業への集中が高くない状態であると言える。このように同じ Bad W.S. であっても、認知・作業状態の時系列データを見ることで、異なる解釈ができる。したがって、例えば前半の区間を作業への集中が十分になるまでの準備期間と解釈して、この時の Bad W.S. を無視し、後半の Bad W.S. のピーク発生時には被験者 s10 に休息を促すことで、適切に疲労を回復させることができると考えられる。

次に、被験者 s22 の 2 桁暗算時の推定結果、図 4.20 について見てみると、被験者 s10 とは異なり、Low W.S. が優位である時間帯がいくつか認められる。さらにそれぞれのピークの後、Optimal W.S. が優位となっている場合が多く、この被験者は適宜休息を取りながら作業を行っていると考えられる。しかし、300 秒を過ぎると、これまで適応度の低かった Bad W.S. のピークが認められるようになり、このピーク出現を皮切りに Optimal W.S. が優位な状態が継続する時間が短くなっており、他の状態のピークが頻繁に出現している。このような状態遷移は、退屈さによる疲労の蓄積が原因であると考えられる。2 桁暗算は単調作業であるため、Optimal W.S. が優位であっても、ある程度の疲労は避けられなかったと思われる。

次に、同被験者の s22 の 4 桁暗算時の推定結果について考察する。図 4.21 を見ると、2 桁加算の時と同様に 300 秒より前では Optimal W.S. が優位であるが、300 秒を過ぎると Low W.S. や Bad W.S. が顕著に現れている。2 桁暗算時の推定結果も考慮すると、被験者 s22 は 300 秒前後から認知・作業状態の現れ方が変化し、Optimal W.S. の出現が抑えられることが分かる。したがって、被験者 s22 には 300 秒付近で休息を促すことで、より良い作業への集中が期待されると考えられる。

以上、4 つの例を用いて推定した認知・作業状態の割合や遷移状況について考察したが、600 秒という短時間の作業でも様々に状態が変化し、被験者によってその状態遷移の様相が異なることが分かった。さらに、それぞれ 4 状態がどのように遷移するかを見ることで、被験者の作業状況のある程度推測でき、適切な休息のタイミングの検討が容易になった。このような考察ができたのも、本研究の提案手法によって各 4 状態の推定結果を時系列解析できたからである。

本研究では、被験者への不快感を与えぬように照明条件や空気質を調節し、実験条件を一定にしていたため、被験者にとってこの実験環境は比較的作業のしやすい環境であった。したがって、本項の考察では執務環境の改善案の検討として、被験者の疲労や休息のタイミングについての考慮に留めているが、今後はこの認知・状態推定手法を

用いて、様々な執務環境の評価や改善案の検討を行うことができると思われる。例えば照明環境を変化させながら執務環境の評価や改善案の検討を行う際は、環境変化の前後で各4状態がそれぞれどのように変動し、どの状態が優位になるのかなどの評価から、執務者にとって作業へ集中しやすい快適な照明環境を提案できると期待される。

4.4.4 今後の課題

本研究で用いたタスクは短期記憶に負荷のかかる暗算タスクで、難易度が2桁と4桁の2種類であったが単調作業であることは否定できない。今後、提案手法を知的作業環境下で用いる際は様々な認知タスクの検討だけでなく、より高度な知的能力を必要とするタスク(航空管制のオペレーションを模したタスクなど)の検討も行わなければならない。

また、本実験では2桁暗算や4桁暗算を被験者にそれぞれ10分間で行わせたが、オフィスワーク等の知的作業環境を考えると比較的短時間である。しかし、長時間のタスク遂行による疲労の蓄積やモチベーション低下が認知・作業状態の変動に影響を与える可能性があるため、今後は長時間の実験も検討し、実際の知的作業環境により近い条件下で提案手法を評価する必要がある。

瞳孔径は認知負荷以外にも、覚醒度^[44]や認知活動量^[38]を反映すると示唆されており、どちらも知的生産性に影響を与える指標である。また、本研究ではECGから心拍数を抽出し認知負荷推定に用いたが、交感・副交感神経を表す心拍変動^[45]はストレスや心的疲労と相関があり、今後の研究ではより正確な知的生産性を評価するために、執務者の内的要因として認知負荷だけでなく他の複数の要因も評価対象にする必要がある。

第 5 章 結論

本研究では、執務者の複数の認知・作業状態を仮定したパフォーマンス-認知モデルを提案し、作業執務者のパフォーマンスと認知負荷、さらに High W.S., Optimal W.S., Bad W.S., Low W.S. の 4 つの認知・作業状態を推定する手法を開発した。パフォーマンス推定では認知タスクの解答時間、分散性及びエラー率、認知負荷の推定では瞳孔径および心拍数を用いた。各推定にはルールベースのファジィ推定を利用し、過去の知見も適用させることでより正確な推定を行った。

提案手法による認知・作業状態推定の妥当性評価として、パフォーマンス及び認知負荷の推定が正当に行われているか確認するため、2種類の難易度のタスクを用いた被験者実験を行った。実験結果から、提案手法において推定されたパフォーマンス及び認知負荷がタスク難易度上昇によってそれぞれ有意に低下及び上昇することが分かったため、提案手法による推定は妥当であると示唆された。さらに、高難易度ほど作業効率の悪い認知・作業状態である Bad W.S. になりやすく、逆に低難易度ほど作業効率の良い Optimal W.S. になりやすいことも提案手法から評価することができた。

また、10分という短時間のタスクであっても被験者の認知・作業状態を時系列で見ると、常に一定ではなく変動しており、被験者によって変動の様子が異なることが分かった。さらに、同じ認知・作業状態であっても、前後の状態遷移から異なる解釈ができ、より詳細に執務者の作業状況を推測できる可能性が示唆された。このように、認知・作業状態を4状態に分類し、提案手法によって時系列的に推定することで、詳細な作業状況の推測が可能となることから、本研究で提案したモデル及び認知・作業状態推定手法の意義は大きく、より効果的な執務環境評価や改善案の検討に貢献できると考えられる。これにより、知的生産性の観点だけでなく執務者の内的要因の観点からも最適とされるような執務環境の実現が期待される。

しかし、本研究の被験者実験で実施したタスクは短期記憶に負荷のかかる暗算タスクであり、多様な知的作業環境下での評価を検討するのならば、より高度、または多様な種類の知的能力を必要とする認知タスク(航空管制のオペレーションを模したタスクなど)の検討も行わなければならない。更に、今後はより長時間のタスク遂行による認知・作業状態の変動も評価し、より詳細な知的生産性評価手法の実現を目指す。

謝 辞

本研究だけでなく様々な知的生産性研究に携わる機会をお与え頂き、学会や研究会等で研究に関する貴重なご意見のみならず、今後社会で活躍するにあたっての様々な助言をして下さった下田 宏 教授に深く感謝いたします。

研究に関して指導して頂いただけでなく、PC 端末や計測デバイス、論文執筆に関して数々の助言をして下さった石井 裕剛 助教に心より感謝いたします。

大学 4 回生から下田研究室に配属されてから 3 年が経ちましたが、自身の研究能力および研究チーム全体の中でプロジェクトを遂行する能力を大いに培うことができたのも、先生方の御指導御鞭撻によるものであると思います。

また共にプロダクティビティチームとして共に知的生産性研究を行い、助言や叱咤激励をして下さった、パナソニック株式会社の大林 史明さん、岩川 幹生さん、エネルギー科学研究科博士 3 回生の宮城 和音さん、昨年度修士課程をご卒業された大石 晃太郎さん、修士 2 回生の内山 皓介くん、修士 1 回生の瀬尾 恭一くん、島村 祐太くん、学部 4 回生の金川 英弘くん心より感謝いたします。

さらに、研究方針の立て方など数多くの助言をして下さり、自身と異なる研究分野であるのにも関わらず、私の研究について様々なご意見をして下さった同研究科博士 2 回生の北村 尊義さんに深く感謝いたします。

気分転換にゲームで一緒に遊んだり、論文執筆に関して助言して頂いたりと研究を進める上で常に支えて頂いた下田研究室の皆様に感謝いたします。

また、研究室生活を送るにあたり、日頃からお世話を頂いた普照 郁美さんに心より感謝いたします。

最後に、4 回生から現在に至るまでの大変有意義な 3 年間を過ごせたのも、様々な方々からの御協力、御助力があつてのことだと思ひます。この場をお借りして、深く感謝の意を表明いたします。

参考文献

- [1] Indoor Health and Productivity Project, Lawrence Berkeley National Laboratory, <http://www.iaqscience.lbl.gov/>, Accessed Feb. 2014.
- [2] 下田宏, 服部瑤子, 富田和宏, 河内美佐, 石井裕剛, 大林史明, 寺野真明, 吉川榮和: オフィスワークのプロダクティビティ改善のための環境制御法の研究- プロダクティビティ評価法 CPTOP の開発-, ヒューマンインターフェースシンポジウム 2006, Vol.1, No.1321, pp.145-150 (2006).
- [3] 岩川幹生, 近辞佑樹, 榎本健治, 宮城和音, 下田宏, 石井裕剛, 寺野真明: 知的生産性改善評価のための新パフォーマンステストの開発, ヒューマンインターフェースシンポジウム 2008, pp.273-280 (2008).
- [4] Wargocki, P., Wyon, D.P., Fanger, P.O.: Productivity is Affected by the AirQuality in Offices, *Healthy Building 2000*, pp.635-640 (2000).
- [5] Thorne, D.R., Genser, S.G., Sing, H.C., Hegge, F.W.: The Walter Reed performance assessment battery, *Neurobehavioral Toxicology & Teratology*, Vol.7, No.4, pp.415-418 (1985).
- [6] Miyagi, K., Kawano, S., Ishii, H., Shimoda, H.: Improvement and Evaluation of Intellectual Productivity Model Based on Work State Transition, *The 2012 IEEE International Conference on Systems, Man, and Cybernetics* pp.1491-1496 (2012)
- [7] Brünken, R., Plass, J.L. , Leutner, D.: Direct Measurement of Cognitive Load in Multimedia Learning, *EDUCATIONAL PSYCHOLOGIST*, Vol.38, No.1, pp.53-61 (2003).
- [8] Paas, F., Tuovinen, J.E., Tabbersm H., Pascal, van Gerven, W. M.: Cognitive Load Measurement as a Means to Advance Cognitive Load Theory, *EDUCATIONAL PSYCHOLOGIST*, Vol.38, No.1, pp.63-71 (2003).

- [9] Paas, F., van Merriënboer, J.J.G.: The efficiency of instructional conditions: An approach to combine mental effort and performance measures, *Human Factors*, Vol.35, No.4, pp.737-743 (1993).
- [10] Ayres, P., Paas, F.: Cognitive Load Theory: New Directions and Challenges, *Applied Cognitive Psychology*, Vol.26, No.6, pp.827-832 (2012).
- [11] Ling, L., Ronnie, T.: Assessing recovery from cognitive load through pen input, *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, pp.1353-1358 (2013).
- [12] Stevens, J.C., Gibert, G., Leung, Y., Zhang, Z.: Evaluating a synthetic talking head using a dual task: Modality effects on speech understanding and cognitive load, *International Journal of Human-Computer Studies*, Vol.71, No.4, pp.440-454 (2013).
- [13] Lomonaco, C., Miller, D.: Environmental Satisfaction, Personal Control and the Positive Correlation to Increased Productivity, Johnson Controls, Inc. (1997).
- [14] 川瀬貴晴: 知的生産性研究の現状と課題 (II ワークプレイスプロダクティビティ評価の実践, <特集> 執務空間の知的生産性・ワークプレイスプロダクティビティ), *建築雑誌*, Vol.122, No.1564, pp.12-13 (2007).
- [15] 村上周三: 知的生産性研究の展望, *空気調和・衛生工学*, Vol.8, No.1, pp.3-8 (2007).
- [16] 村上周三: 建築空間と知的活動の階層モデル, 2012.3.15 知的生産性委員会, 資料 No.5 (2012).
- [17] Hess, E.H., Polt, J.M.: Pupil size in relation to mental activity during simple problem-solving, *Science*, Vol.143, pp.1190-1192 (1964).
- [18] Stanners, R.F., Coulter, M., Sweet, A.W., Murphy, P.: The Pupillary Response as an Indicator of Arousal and Cognition, *Motivation and Emotion*, Vol.3, No.4, pp.319-340 (1979).
- [19] 山本博巳, 清水康敬: 数字列記憶実験における瞳孔面積と瞬目時間の変化, *テレビジョン学会誌*, Vol.45, No.3, pp.370-376 (1991).

- [20] Halverson, T., Estep, J., Christensen, J., Monnin, J.: Classifying Workload with Eye Movements in a Complex Task, Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol.56, No.1, pp.168-172 (2012).
- [21] Iqbal, S. T., Adamczyk, P. D., Zheng, X. S., Bailey, B. P.: Toward an Index of Opportunity: Understanding Changes in Mental Workload during Task Execution, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp.311-320 (2005).
- [22] 石橋富和, 大谷璋, 三浦武夫: 精神負担の指標としての心拍数, 産業医学, Vol.10, No.7, pp.377-379 (1968).
- [23] Turner, J.R., Carroll, D.: Heart rate and oxygen consumption during mental arithmetic, a video game, and graded exercise: further evidence for metabolically exaggerated cardiac adjustment, Psychophysiology, Vol.22, No.3, pp.261-267 (1985).
- [24] Douglas Carroll, J.Rick Turner, Rekha Prasad: The effects of level of difficulty of mental arithmetic challenge on heart rate and oxygen consumption, International Journal of Psychophysiology, Vol.4, No.3, pp.167-173 (1986).
- [25] G. Mulder, L. J. M. Mulder: Information processing and cardiovascular control, Psychophysiology, Vol.18, No.4, pp.392-402 (1981).
- [26] Backs, R.W., Seljos, K.A.: Metabolic and cardiorespiratory measures of mental effort: the effects of level of difficulty in a working memory task, International Journal of Psychophysiology, Vol.16, No.1, pp.57-68 (1994).
- [27] Kennedy, D.O., Scholey, A.B.: Glucose administration, heart rate and cognitive performance: effects of increasing mental effort, Psychopharmacology, Vol.149, No.1, pp.63-71 (2000).
- [28] Schumacher, K.J.: ファジィ集合 自然言語演算とリスク解析, 啓学出版, 第1章-第3章 (1990).
- [29] 長町三生: ファジィ化製品開発の基礎と実際-ニューロ・ファジィ理論から活用事例まで, 海文堂出版, 第1章 (1991).

- [30] MATLAB ,MathWorks, Inc., <http://www.mathworks.co.jp/products/matlab/>, Accessed Feb. 2014.
- [31] Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information, *Psychological Review*, Vol.63, No.2, pp.81-97, (1956).
- [32] Card, S.K., Moran, T.P., Newell, A.: *The Psychology of Human-Computer Interaction*, Erlbaum Associates (1983).
- [33] 河野 翔: 意識的な休息に着目した知的生産性変動モデルの提案と評価, 京都大学 エネルギー科学研究科修士論文 (2012).
- [34] Suzuki, S., Matsui, T., Kawahara, H., Ichiki, H., Shimizu, J., Kondo, Y., Gotoh, S., Yura, H., Takase, B., Ishihara, M.: A non-contact vital sign monitoring system for ambulances using dual-frequency microwave radars, *Medical & Biological Engineering & Computing*, Vol.47, No.1, pp.101-105 (2009).
- [35] Cardiio, Cardiio, Inc., <http://www.cardiio.com/>, Accessed Feb. 2014.
- [36] facelab, Seeing Machines, Inc., <http://www.seeingmachines.com/product/facelab/>, Accessed Feb. 2014.
- [37] Davson, H.: *The eye*, Academic Press, Vol.3, pp.255-337 (1969).
- [38] Marshall, S.: *Method and Apparatus for Eye Tracking and Monitoring Pupil Dilation to Evaluate Cognitive Activity*, Patent 6,090,051, USA, (2000).
- [39] 松永勝也: 瞳孔運動の心理学, ナカニシヤ出版, 第2章 pp.78-81 (1990).
- [40] Bjerregaard, P.: Mean 24 hour heart rate, minimal heart rate and pauses in healthy subjects 40-79 years of age; *European Heart Journal*, Vol.4, No.1, pp.44-51 (1983).
- [41] Raskin, C., David, Kotses, H., Bever, James: Cephalic Vasomotor and Heart Rate Measures of Orienting and Defensive Reflexes, *Psychophysiology*, Vol.6, pp.149-159 (1969).

- [42] Hart S. G., Staveland L. E.: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research, Hancock P. A., Meshkati N. (Eds.), Human mental workload, pp.139-183, (1988).
- [43] Franke RH, Kaul JD: The Hawthorne experiments: First statistical interpretation, American Sociological Review, Vol.43, pp.623-643 (1978).
- [44] 西山潤平, 谷田公二, 楠見昌司, 平田 豊: 瞳孔揺らぎを指標とした覚醒状態評価, 生体医工学, Vol.46, No.2, pp.212-217, (2008).
- [45] Camm, A., Malik, M., Bigger, J., Breithardt, G., Cerutti, S., Cohen, R., Coumel, P., Fallen, E., Kennedy, H.L., Kleiger, R.E., Lombardi, F., Malliani, A., Moss, A.J., Rottman, J.N., Schmidt, G., Schwartz, P.J., Singer, D.: Heart rate variability: Standards of measurement, physiological interpretation, and clinical use; Circulation, Vol.93, No.5, pp.1043-1065 (1996).