

令和 6 年 5 月 28 日現在

機関番号：12608
研究種目：基盤研究(B)（一般）
研究期間：2020～2022
課題番号：20H01718
研究課題名（和文）プログラムコード読解作成能力育成のための認知行動に基づく形成的評価手法の確立

研究課題名（英文）Development of formative evaluation method based on cognitive behavior for evaluation of program code reading comprehension ability

研究代表者
中山 実（Nakayama, Minoru）
東京工業大学・工学院・教授

研究者番号：40221460
交付決定額（研究期間全体）：（直接経費） 7,200,000円

研究成果の概要（和文）：プログラムコードの読解と作成の能力育成のために形成的な評価手法を確立し、効果的な指導法を開発することを目的とした。読み行動としての眼球運動におけるサッカードや瞳孔径の変化を基に、読者のプログラムコード読解力を調べた。特注量を基に読む人の属性や能力との関連を調べた。コードの読み行動の注視順序性から能力評価が行えること、モデリングの手法から読み段階での注意レベルを評価することで能力や状態評価ができることを示した。プログラムコード作成能力を学習者相互で評価するための相互評価手法を開発し、大学でのプログラム作成学習での相互評価で教員の評価に準じた評価ができることを示した。

研究成果の学術的意義や社会的意義
文書の読み行動を眼球運動で検討する研究を、コマンド記述で表現されるプログラムコードに応用し、能力評価が難しいプログラムの読解作成能力評価を検討した。読み行動を単語単位の注視行動でモデル化する手法が検討されており、事前に検討された属性を用いることでプログラミング能力の推定ができることを示した。また、読み過程での注意レベルの推定もできることを示した。プログラミング能力の評価には、多様な観点が必要であることから、学習者同士による相互評価も有効である。そこで、その評価法の妥当性確立のために、学習者相互評価から教員評価に相当する評価値を推定する手法を開発した。

研究成果の概要（英文）：Formative evaluation methodologies for programming ability were established in order to develop reading and writing codes and these appropriate instructions. Eye movements as reading behavior including saccades and pupil reactions were introduced to estimate ability of program code comprehension and levels of expertisness. The results showed a possibility to predict performance and temporal conditions such as level of attention during reading with modelling technique.

Another evaluation procedure of the ability was developed using peer assessment within learners, survey results at a university class provided a comparable performance ratings with instructors.

研究分野：教育工学

キーワード：プログラムコード 読解作成能力 認知行動 形成的評価

1. 研究開始当初の背景

プログラミング学習においては、学習者の能力を適切に評価・指導することが重要である。特にプログラミング能力は、多様な評価基準で検討することが求められることから容易ではない。本研究では、プログラムコードの読解過程を調べることによる評価可能性と、協調学習の観点での学習者同士による相互評価を基にした能力推定の可能性を検討した。

それぞれの先行研究の概要を以下に示す。

- (1) プログラムコードの読解における眼球運動の特徴は、プログラミング能力が反映する[1,2]。そこで、この観察における眼球運動を計測によって学習者の状態を把握して、指導や支援を適切に行う方法が検討されている。この検討をおこなうために、EMIP(Eye Movement in Programing)のワークショップでさまざまな熟達度のプログラマの眼球運動を計測・共有してこの課題を検討している[3,4]。
- (2) 大学でのプログラミング指導の検討事例において、能力評価と指導の方法として、学習者同士による相互評価を導入された[8-10]。単純な相互評価の集計でも学習指導において効果があるが、評価精度を向上させるために、項目反応理論(Item Response Theory)に基づいた評価が期待されている[11]。

2. 研究の目的

先行研究を基に以下の内容について検討した。

- (1) プログラムコード読解での眼球運動による能力評価については、用いる眼球運動特徴量や、評価される能力そのものを明確にする必要がある。公開されているデータを用いて、複数の分析手法による評価可能性と指導への利用について検討した。
- (2) プログラミング能力を、学習者による評価によってより正確に計測するための手法や評価の条件を明らかにし、評価条件による利用可能性について検討した。

3. 研究の方法

(1) プログラムコード読解での眼球運動による能力評価：EMIP データセットを用いて以下の分析を行った。データセットは世界各地のボランティア 269 人の参画で生成された。難易度の異なる 2 種の Java コードを実験刺激として、ランダムな順番で提示され、コードを読んでいる時間内と用意されたコードに関する質問に回答する時間内の眼球運動が計測された。眼球運動のサンプリングレートは 250Hz であった。プログラムの内容は以下の 2 つである[3,4]。

- Vehicle プログラム：コンストラクタ、加速メソッドとメイン関数が記述されており、メイン関数内でクラスを作成し加速を計算する。

- Rectangle プログラム：コンストラクタ、面積計算メソッドとメイン関数が記述されており、メイン関数内でクラスを作成しオブジェクトの面積を計算する。

それぞれのコードには内容に関する質問が 1 問だけ用意され、それぞれの回答正誤が記録された。この他、プログラミング経験年数や自己評価など、参加者の属性が評定値として記録されていた。分析には、欠損データを除いた 216 人で構成された。

プログラミング能力を、質問項目や属性回答を基に因子分析し、能力値を推定した。これと眼球運動の特性値との合成量で評価できるかを検討した。用いた眼球運動特徴量は、眼球運動特徴量 6 種、垂直方向周波数パワー12 種、水平方向周波数パワー12 種、注視遷移確率 3 種、ブロック内の注視、スキャンパス情報 36 種の計 69 特徴量を用いた。特徴量が多いことから、これらも因子分析によって 6 因子に集約して評価した。各因子の特徴量は以下の通り。

- 因子 1 は水平及び垂直方向の注視位置変動
- 因子 2 は読解全体の眼球運動特徴
- 因子 3 は main 関数における注視特徴
- 因子 4 は処理関数における平均化された注視特徴
- 因子 5 は処理関数における注視特徴
- 因子 6 はコンストラクタにおける読解初期の注視特徴

これらの関係をガウスカーネルによるサポートベクトル回帰によって推定可能性を評価した。

(2) プログラムコードの学習者相互による評価点の収集
イタリアの大学でのプログラミング言語 C の実習授業で、2 年度で相互評価を実験的に行った[12,13]。内容はプログラミング基礎で、変数、ブール代数に加えて、入出力ライブラリ関数、分岐、配列などであった。授業の参加者は、2 回とも 60 名程度の学生であった。相互評価は 4～

```
public class Vehicle {
    String producer, type;
    int topSpeed, currentSpeed;

    public Vehicle (String p, String t, int tp) {
        this.producer = p;
        this.type = t;
        this.topSpeed = tp;
        this.currentSpeed = 0;
    }

    public int accelerate (int kmh) {
        if ((this.currentSpeed + kmh) > this.topSpeed) {
            this.currentSpeed = this.topSpeed;
        } else {
            this.currentSpeed = this.currentSpeed + kmh;
        }
        return this.currentSpeed;
    }

    public static void main (String args []) {
        Vehicle v = new Vehicle ("Audi", "A6", 200);
        v.accelerate (10);
    }
}
```

図 1. 提示コードとブロック分割例

5 回実施された。1~4 回は対面で、5 回目はオンラインで実施された。それぞれの学生は 3 人の peer の課題回答を 100 点満点で評価した。ただし、収集されたデータでは、毎回、全員の回答を得ることはなかった。なお、教員も個別に評価点を与えた。また、中間および最終成績も別途与えた。

相互評価の評価点の分析には IRT を用いた。プログラミング能力のようなパフォーマンス評価では、以下のような評価者の特性が能力評価の精度に影響する[14,15]。

- 厳しさ：パフォーマンスに対して全体的に低い評点を与える傾向。
- 一貫性：評価基準が一貫している程度を表し、評点が安定している。
- 尺度範囲の制限：特定の評点に集中する傾向を表す。

評価項目の特性としては以下のパラメータが設定される。

- 困難度：得られる評点が全体として低くなる特性。
- 識別力：能力がその課題に対するパフォーマンスに反映される度合。

これらのパラメータを考慮した IRT モデルである一般化多層ラッシュモデル(GMFRM: the generalized many-facet Rasch model)[14,15]を適用した。

4. 研究成果

(1) プログラミング能力の眼球特徴因子による回帰推定：

回帰分析による推定値と熟練度との関係を図 2 に示す。回帰での因子の寄与を検討した。コード 2 種に共通して確認された特徴は、因子 1 すなわち水平及び垂直方向の注視位置変動は熟練度が高い読解者ほど特徴量が増加し、因子 6 すなわちコンストラクタにおける読解初期の注視特徴は熟練度が高い読解者ほど特徴量が減少することが分かった。

視線移動成分とは、水平、垂直方向の視線位置変動を時系列データと捉えた際の周波数パワーであり、因子分析による高い因子負荷が見られたのは 1.5-6.0Hz の周波数帯であった。特に 4.5-6Hz の周波数帯が顕著な特徴として集約されたため、高周波領域のパワーが要因と大きく関係していると考えられる。高周波領域のパワーが高い要因としてサッカードの発生が考えられる。コード内各行の要素を順番に読んでいく場合、注視点間距離は短く、発生するサッカードは高速にならない。しかし、行内の要素において文頭から文中を飛ばして読み、文末の必要な箇所だけ読む読解方法を行なった場合、注視点間距離が長く、高速なサッカードが発生する。注視点間の移動をスキャンパスとして分析したが、速度の閾値を設定することによってサッカード検出して、詳細な変化が検討できる。

また、熟練度が高い読解者ほど、コンストラクタにおける読解初期の注視特徴の因子得点が小さくなった。これは読解方略の影響と考えることができた。読解初期はスキャンと呼ばれるコード全体を上から下に眺める動作が知られており[5]、注意深くスキャンすることで、ブロック内の各箇所を細かく読み、細かな要素それぞれへの注視持続時間が短く、平均スキャンパス長やスキャンパスの持続時間が短くなったと考えられる。

なお、これらの検討を基に時系列的な眼球運動変化を用いた推定も検討した。コード表示画面内での眼球運動に着目した場合での回答正誤を高精度に推定できた[6]。また、プログラムコードの行内での読み行動を、コード構成要素をトークンとして、トークン間の推移に着目することで、能力評価ができることを確認した[7]。

(2) 学習者の相互評価点から個別の能力推定
相互評価で得られた測定データ間の関係として、Peer から与えられた評価点の平均点と教員評価点との関係を図 3 に示す。図中には授業 1 と授業 2 の結果を示している。授業 2 では教員の評価点の一部の学生で低くなっており、授業 1 でのデータ散布の傾向が異なっている。

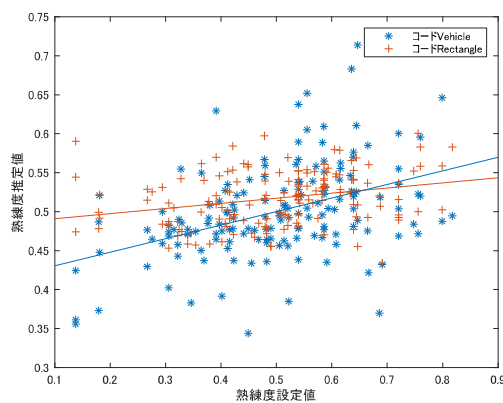


図 2. 熟練度の設定値とサポートベクトル回帰による予測値との散布図

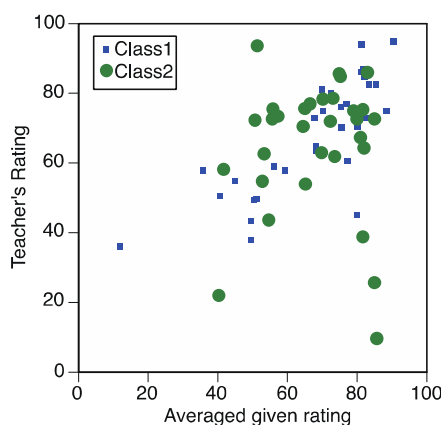


図 3. 相互評価の平均点と教員評価点

相関関係を調べたところ、授業 1 では正の相関であった ($r=0.79$) が、授業 2 では無相関であった ($r=-.05$)。授業 2 では、教員の評価点が低い学生でも、Peer が高い評価を与えているために、無相関が示された。教員評価点の分散は 2 つの授業間では同程度 (授業 1:SD=16.3; 授業 2:SD=16.1) で一貫性があると考えられる。Peer の評価について検討が必要である。

教員の最終評価点との関係を調べた。相互評価の平均点は教員の最終評価点が低いほど分散が大きくなる傾向が見られ、特に授業 1 (SD=18.2) でのばらつきが授業 2 (SD=13.4) よりも大きい。教員の最終評価点の分布は、2 つの授業ではほぼ同程度である。このため、相関係数を調べると授業 1 では正の相関が見られる ($r=0.47$) が、授業 2 では相関係数は小さかった ($r=0.10$)。

推定結果と授業での評価点との関係を調べた。相互評価点に前述の手法を適用して得られた、推定された能力値との関係を調べた。その結果を図 4 に示す。授業 1 では教員の中間評価と能力推定値は強い相関関係が見られたが、授業 2 では低かった。教員の最終評価点と能力推定値との相関関係を調べると、授業 1 ($r=0.48$)、授業 2 ($r=0.22$) で、授業 2 での関連が改善された。

これらの相関関係によって相互評価による評価の妥当性とは判断できないが、IRT に基づく推定によって、相互評価された学習者の能力評価に改善が見られ、評価の情報としての利用可能性が確認できた。

能力値推定において授業 1 と授業 2 での相互評価参加者の特性には違いが見られたので、その特性を分析した。授業間での違いなどを検出するために、授業ごとに算出した能力値、一貫性、厳密さの特性値を用いてクラスタ分析した。その結果、4 クラスタが抽出された。特性値のうち、能力値と厳密さの分散が相対的に大きいので、この 2 変数の平面で 4 つのクラスタを表現して図 5 に示す。図中には授業の違いを明示しているが、クラスタへの偏りは授業 1 の外れ値で構成される Clust2 以外には見られない。能力が高い Clust1 と低い Clust3 は厳密さが高く、能力が中程度の Clust4 では厳密さが低くなった。

この結果は、授業間での違いが小さいことから、相互評価の評定における共通性を示している。能力値が中程度の評定者はやや曖昧な評定をする傾向があるため厳密さが低く、能力値が高い、あるいは低い場合は、厳密に評定している。これらは学習動機のような授業への参加態度による影響も考えられるため、より具体的な因果関係については、別途の調査が必要になる。相互評価の評定活動の特性から、学習態度属性が推定できるのであれば、個別の学習指導への活用も期待される。特に、プログラミング授業のような個別にプログラムを作成する実習形式での評価であることから、評定の属性を用いた指導可能性が期待できる。

参考文献：

- [1] T. Busjahn, R. Bednarik, A. Begel, M. Crosby, J.H. Paterson, C. Schulte, B. Sharif, S.Tamm, "Eye movements in code reading: Relaxing the linear order" 2015 IEEE 23rd International Conference on Program Comprehension, pp.255–265, 2015
- [2] R. Bednarik, L. Budde, B. Heinemann, C. Schulte, H. Vrzakova, "Eye-movement Modeling Examples in Source Code Comprehension: A Classroom Study" In 18th Koli Calling International Conference on Computing Education Research (Koli Calling '18), 22-25.,2018
- [3] R. Bednarik, et al. "EMIP: The eye movements in programming dataset" Technical Report. 2018.
- [4] K. Martin, "Data Flow Metrics in Program Comprehension Tasks" EMIP '18 Proceedings of the Workshop on Eye Movements in Programming, Article No.2, pp.1–6, 2018
- [5] H. Uwano, M. Nakamura, A. Monden, K. Matsumoto, "Analyzing individual performance of source code review using reviewers' eye movement" Proceedings of the 2006 symposium on Eye tracking research & applications pp133–140, 2006

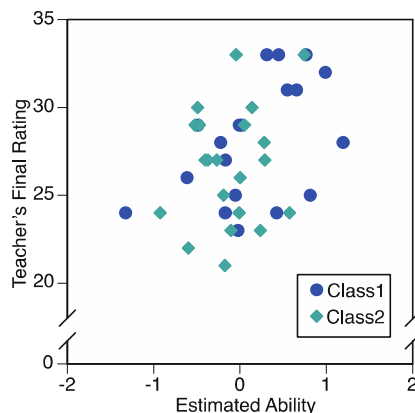


図 4. 推定した能力値と教員評価点

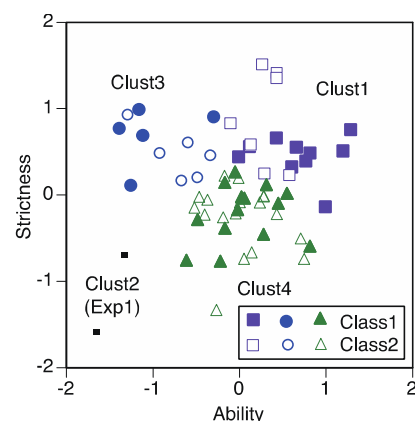


図 5. IRT 特性値のクラスタ分析の結果

- [6] Y. Yan, M. Nakayama, "Predicting code reading test answers by using eye movement features", 信学技報, ET2021-11, vol.121, no.174, 17-22, 2021
- [7] M. Sun, M. Nakayama, "Program code navigation model for individuals based on LSTM with co-clustering", ETRA-EMIP, 88, 1-7, 2023.
- [8] A.-K. Amal, K. and D. Marie, "Evaluating a peer assessment approach in introductory programming course," in Proc. UKICER '20, 2020, pp. 51–58.
- [9] K. Georgouli and C. Sgouropoulou, "Collaborative peer-evaluation learning results in higher education programming-based courses," in Proc. ICBL2013 Int. Conf. on Interactive Computer aided Blended Learning, 2013, pp. 309–314.
- [10] X.-M. Wang, G.-J. Hwang, Z.-Y. Liang, and H.-Y. Wang, "Enhancing students' computer programming performances, critical thinking awareness and attitude towards programming: An online peer-assessment attempt," Educational Technology & Society, vol. 20, pp. 58–68, 2017.
- [11] M. Uto and M. Ueno, "Empirical comparison of item response theory models with rater's parameters," Heliyon, vol. 4, pp. 1–32, 2018.
- [12] M. D. Marsico, A. Sterbini, and M. Temperini, "Towards a quantitative evaluation of the relationship between the domain knowledge and the ability to assess peer work," in Proc. of iTHET 2015, Lisbon, Portugal, 2015, pp.1–6.
- [13] M. D. Marsico, A. Sterbini, and M. Temperini, "Modelling peer assessment as a personalized predictor of teacher's grades: the case of OpenAnswer," in Proc. of iTHET 2016, Istanbul, Turkey, 2016, pp. 1–5.
- [14] M. Uto and M. Ueno, "Item response theory without restriction of equal interval scale for rater's score," in Proc. International Conference on Artificial Intelligence in Education, 2018, pp. 363–368.
- [15] M. Uto and M. Ueno, "A generalized many-facet Rasch model and its Bayesian estimation using Hamiltonian Monte Carlo," Behaviormetrika, vol. 47, pp. 469–496, 2020.

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 6件/うち国際共著 4件/うちオープンアクセス 5件）

1. 著者名 Minoru Nakayama, Filippo Sciarrone, Marco Temperini, Masaki Uto.	4. 巻 20
2. 論文標題 An Item Response Theory Approach to Enhance Peer Assessment Effectiveness in Massive Open Online Course	5. 発行年 2022年
3. 雑誌名 Int. J. Distance Education Technologies	6. 最初と最後の頁 1-19
掲載論文のDOI（デジタルオブジェクト識別子） 10.4018/IJDET.313639	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Wioletta Nowak, Minoru Nakayama, Tomasz Krecicki, Andrzej Hachol	4. 巻 6(23)
2. 論文標題 Detection procedures for patients of Alzheimer's disease using waveform features of pupil light reflex in response to chromatic stimuli	5. 発行年 2020年
3. 雑誌名 EAI Endorsed Transactions on Pervasive Health and Technology	6. 最初と最後の頁 1-11
掲載論文のDOI（デジタルオブジェクト識別子） 10.4108/eai.17-12-2020.167656	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Akiko Teranishi, Minoru Nakayama, Theodor Wyeld, Mohamad A. Eid	4. 巻 29(2)
2. 論文標題 Evaluation of online team-based game development using SNS tools	5. 発行年 2021年
3. 雑誌名 International Journal of Innovation and Learning	6. 最初と最後の頁 181-206
掲載論文のDOI（デジタルオブジェクト識別子） 10.1504/IJIL.2021.112995	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
1. 著者名 Minoru Nakayama, Yoshiya Hayakawa	4. 巻 7(23)
2. 論文標題 Influence of Task-evoked Mental Workloads on Oculo-motor indices and their connections	5. 発行年 2021年
3. 雑誌名 EAI Trans. Context-aware Systems and Application	6. 最初と最後の頁 1-10
掲載論文のDOI（デジタルオブジェクト識別子） 10.4108/eai.4-2-2021.168649	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 篠田瑞生, 中山実, 伊藤泉	4. 巻 23(1)
2. 論文標題 サッカーの経験による試合映像への注視行動に関する検討	5. 発行年 2021年
3. 雑誌名 ヒューマンインターフェース学会論文誌	6. 最初と最後の頁 47-50
掲載論文のDOI (デジタルオブジェクト識別子) 10.11184/his.23.1_47	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Minoru Nakayama, Wioletta Nowak, Anna Zarowska	4. 巻 LNBIP471
2. 論文標題 Symptoms of Dementia in Elderly Persons Using Waveform Features of Pupil Light Reflex	5. 発行年 2023年
3. 雑誌名 Information Technology for Management: Approaches to Improving Business and Society	6. 最初と最後の頁 91-107
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-031-29570-6_5	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

[学会発表] 計13件 (うち招待講演 0件 / うち国際学会 13件)

1. 発表者名 Tomomi Okano, Minoru Nakayama
2. 発表標題 Research on Time Series Evaluation of Cognitive Load Factors using Features of Eye Movement
3. 学会等名 ACM ETRA 2022 / COGAIN workshop 2022 (国際学会)
4. 発表年 2022年

1. 発表者名 Minoru Nakayama
2. 発表標題 Predicting individual sentiment for emotion-evoking pictures using metrics of oculo-motors,
3. 学会等名 Invormation Visualisation (国際学会)
4. 発表年 2022年

1. 発表者名 Hiroto Harada, Minoru Nakayama
2. 発表標題 Estimation of reading ability of program codes using features of eye movements
3. 学会等名 ACM ETRA 2021 / EMIP workshop 2021 (国際学会)
4. 発表年 2021年

1. 発表者名 Yudai Niitsu, Minoru Nakayama
2. 発表標題 Effects of measurement time and presentation size conditions on biometric identification using eye movements
3. 学会等名 ACM ETRA 2021 / EyeSec workshop 2021 (国際学会)
4. 発表年 2021年

1. 発表者名 Minoru Nakayama, Masaki Uto, Marco Temperini, Filippo Sciarrone
2. 発表標題 Estimating Ability of Programming Skills using IRT based Peer Assessments
3. 学会等名 ITHET / IEETel 2021 (国際学会)
4. 発表年 2021年

1. 発表者名 Hiroto Harada, Minoru Nakayama
2. 発表標題 Eye movement features in response to comprehension performance during the reading of programs
3. 学会等名 ACM ETRA2020, EMIP-WS (国際学会)
4. 発表年 2020年

1. 発表者名 Minoru Nakayama, Filippo Sciarrone, Masaki Uto, Marco Temperini
2. 発表標題 Estimating Student ' s Performance Based on Item Response Theory in a MOOC Environment with Peer Assessment
3. 学会等名 Methodologies and Intelligent Systems for Technology Enhanced Learning (Mis4TEL) (国際学会)
4. 発表年 2020年

1. 発表者名 Minoru Nakayama, Filippo Sciarrone, Masaki Uto, Marco Temperini
2. 発表標題 Impact of the number of peers on a mutual assessment as learner ' s performance in a simulated MOOC environment using the IRT model
3. 学会等名 Information Visualisation2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Taijiro Shiraishi, Minoru Nakayama
2. 発表標題 Deviations of eye movements and head rotation during response tasks using targets and HMD
3. 学会等名 Information Visualisation2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Minoru Nakayama, Wioletta Nowak, Anna Zarowska
2. 発表標題 Prediction Procedure for Dementia Levels based on Waveform Features of Binocular Pupil Light Reflex
3. 学会等名 ACM ETRA2023 (国際学会)
4. 発表年 2023年

1. 発表者名 Ming Sun, Minoru Nakayama
2. 発表標題 Program Code Navigation Model for Individuals Based on LSTM with Co-Clustering
3. 学会等名 ACM ETRA 2023 / EMIP 2023 (国際学会)
4. 発表年 2023年

1. 発表者名 Mateusz Dubiel, Minoru Nakayama, Xin Wang
2. 発表標題 Modelling Attention Levels with Ocular Responses in a Speech-in-Noise Recall Task
3. 学会等名 ACM ETRA 2023 / COGAIN 2023 (国際学会)
4. 発表年 2023年

1. 発表者名 Minoru Nakayama, Takahiro Ueno
2. 発表標題 Estimation of Latent Attention Resources Using Microsaccade Frequency During a Dual Task
3. 学会等名 ACM ETRA 2023 (国際学会)
4. 発表年 2023年

〔図書〕 計1件

1. 著者名 Minoru Nakayama	4. 発行年 2021年
2. 出版社 Springer Singapore	5. 総ページ数 106
3. 書名 Pupil Reactions in Response to Human Mental Activity	

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分 担 者	宇都 雅輝 (Uto Masaki) (10732571)	電気通信大学・大学院情報理工学研究科・准教授 (12612)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関