

敵対的生成ネットワークを用いた 作業者の性格ごとの頭部・眼球運動の特徴可視化

山下 純平[†] 蟹江 大成^{††} 瀧本 祥章^{†††} 小矢 英毅[†] 片岡 明[†]

大石 晴夫[†] 熊田 孝恒^{††}

[†] 日本電信電話株式会社 NTT アクセスサービスシステム研究所 〒239-0847 神奈川県横須賀市光の丘 1-1

^{††} 京都大学大学院 情報学研究科 〒606-8501 京都府京都市左京区吉田本町

^{†††} 日本電信電話株式会社 NTT サービスエボリューション研究所 〒239-0847 神奈川県横須賀市光の丘 1-1

E-mail: [†], ^{††} {junpei.yamashita.vd,yoshiaki.takimoto.ar,hidetaka.koya.gk,haruo.oishi.nw}@hco.ntt.co.jp

^{†††} {kanie.taisei.83r@st, t.kumada@i}.kyoto-u.ac.jp

あらまし 本研究では、半教師あり相互情報量最大化敵対的生成ネットワークを用いて、作業者の性格と、オフィス作業を行う際の頭部・眼球運動の特徴の関係を可視化できる手法を提案する。既存研究では、眼球運動の特徴から性格を推定するブラックボックスなモデルを用いているため、両者間の関係性を可視化することはできない課題があった。加えて、現実の作業を対象とするには、1) 複雑な特徴抽出をモデルが自動で行うこと、2) 個人ごとの計測ノイズを分離できること、の2つの要件を満たす必要がある。提案手法は、これらの要件を満たしているため、現実の作業を対象とした分析が可能である。評価実験の結果、提案手法により、作業の内容が転記・会話課題として統制して与えられた状況で、未知の個人について、開放性と呼ばれる性格特性を推定できるモデルが得られた。得られたモデルを用いて、開放性の強さごとに頭部・眼球運動データを生成して特徴を分析した結果、開放性が高い場合には、頭部・眼球運動の周期が大きくなることが示唆された。これは、心理学研究において、開放性が高い場合に生じると報告される特徴と一致しており、分析結果の妥当性が支持されるものであった。

キーワード 頭部運動, 眼球運動, 性格, Big Five, 敵対的生成ネットワーク, GAN

1 序 論

個人の性格と特徴的な行動の関係を明らかにすることで、行動から心理状態を推定する技術の実現を目指している。例えば、知的好奇心が高い性格の個人は、抽象的な画像を呈示されると、画像の意味を考えながら行動するパターンを示すと報告されている [1]。このような関係が現実でも再現するのであれば、行動の特徴から、現在取り組んでいる作業に好奇心を持っているかなどの心理状態を取得でき、作業者のモチベーションに沿った業務割り振りなどに活用できる可能性がある。

本研究では、上記の応用を見据え、オフィス作業を対象として、性格と特徴的な行動の関係を分析した。オフィス作業は、目立った身体運動を伴わない認知的行動（PC 作業、打合せなど）が中心である。このような作業においては、視覚による情報収集が重要な役割を果たしている [2]。そこで、本研究では、視線移動のパターンを行動の特徴として分析した。

性格と視線移動の関係を示唆する機械学習を用いた既存研究には、Berkovsky ら [3]、Hoppe ら [4] がある。しかしながら、これらの研究では、視線移動の特徴から性格を推定するブラックボックスなモデルを用いているため、視線移動の特徴が、どのようにして性格の推定に寄与しているのかを具体的には可視化できない課題がある。さらに、現実の作業を対象とした分析

を行うためには、モデルが次の2つの要件を満たしている必要がある。まず、1) 多様な行動が現れる現実の作業においては、性格と関係する特定の視線移動の特徴を事前に想定することが難しいため、人手による特徴抽出なしでモデルを構築できることが必要である。次に、2) 計測ノイズが大きく、その個人差も大きくなる現実の作業を対象とするには、性格の違いに結びついた特徴と、個人ごとの計測ノイズに結びついた特徴を切り分けられるモデルが必要である。

そこで、本研究では、性格と視線移動の特徴の関係を可視化する、2つの要件を満たした手法として、半教師あり相互情報量最大化敵対的生成ネットワーク（semi-supervised Information Maximizing Generative Adversarial Networks: ss-InfoGAN）を用いた手法を提案する。提案手法は、視線移動の特徴から性格ラベルを推定する過程と、性格ラベルに対応した低次元の潜在変数から多次元の視線移動データが生成される過程を同時にモデル化する。これにより、視線移動の特徴から性格を推定するだけでなく、その逆に、ある性格に典型的な頭部・眼球運動を生成することが可能となる。すなわち、ある性格の強さを変えた場合に、視線移動のパターンがどのように変化するかを直接可視化することが可能である。それゆえ、本研究の目的である性格と特徴的な行動の関係の分析が可能となる。また、提案手法は、1) 人手による特徴抽出が不要であり、2) 性格ラベルに結びついた視線移動の特徴と、個人ラベルに結びついた特徴

(個人ごとの計測ノイズ)を分離することが可能である。これにより、現実の作業を対象とした分析が可能となっている。

2 関連研究

2.1 性格特性とその検査

2.1.1 性格特性

人間の性格を記述するにあたって、特性論という立場では、人間が共通のパーソナリティ構造(特性)をもつと仮定し、個人の性格を共通特性の値の強さの組み合わせとしてとらえる [5]。提案されている特性には様々なものがあるが、Big Five [5] と呼ばれる 5 つの共通特性が最もよく用いられている [6]。Big Five では、(経験への) 開放性 (Openness)、誠実性 (Conscientiousness)、外向性 (Extraversion)、協調性 (Agreeableness)、神経症傾向 (Neuroticism) の 5 つのパーソナリティ特性から個人の性格を記述する。開放性は好奇心・知性などの側面、誠実性は自己規律・組織的である程度・意欲的である程度などの側面、外向性は、人の社会的な外向性の程度、協調性は好意的である程度・温厚・同情的である程度などの側面、神経症傾向は感受性・心配性・不安性などの側面を表す。本研究では、Big Five を性格特性として用いている。

2.1.2 質問紙法による性格検査

一般に、個人の性格を取得するには、質問紙法が用いられる [5]。質問紙法とは、性格を推定したい対象者に、性格を反映していると考えられる単語によって構成される質問に回答してもらう手法である。例えば、「温和だと思う」「同調的だと思う」などの質問に自身がどれほど当てはまるかを評定してもらい、これらの評定値から、「協調性の高さ」などの値が算出される。McCrae ら [5] は、Big Five を測定するための質問紙である Neuroticism-Extraversion-Openness Personality Inventory (NEO-PI) を開発して他のテストとの妥当性を検証し、安定して 5 因子が抽出されることを示した。しかし、NEO-PI は英語圏で確立された質問紙であるため、日本で検査を実施するには日本語版の質問紙が必要となる。また、NEO-PI の質問数は非常に多いため、実施に時間がかかる難点がある。そこで、Iwai ら [6] は、自然言語処理の手法を用いて、Big Five の各パーソナリティ特性に関連している日本語の単語を抽出し、心理学的な実験を通して、測定に最も有効であった少数の単語を選定し、計 20 の質問のみで Big Five を測定する Trait Descriptors Personality Inventory (TDPI) を提案している。本研究では、性格特性の取得に TDPI を用いている。

2.2 視線計測と視線移動の特徴による性格推定

2.2.1 視線計測

視線移動は、主として眼球が運動することと、眼球が位置する頭部が運動することによって実現される。このうち、眼球運動の取得においては、赤外線カメラによる専用装置を用いた計測方法が広く知られている。しかし、この方法では、高精度に眼球運動が取得できるが、大容量データを高速で処理する必要性から、日常的に装着して計測を行うことは容易ではない [7]。

一方で、カメラを用いずに眼球運動を計測する方法として、眼電位 (Electrooculography: EOG) を用いた方法がある。EOG では、眼球が前方に正、後方に負の電位を持っていることを利用して、眼球の周囲に電極を貼ることにより、眼球の相対的運動を推定できる [7]。既存研究において、Ishimaru ら [7] は、通常の眼鏡と変わらないサイズ・見た目の、眼電位計を備えたメガネ型ウェアラブルデバイス JINS MEME によって、広い場面で眼球運動を計測可能であることを示している。

頭部運動は、小型・軽量の慣性センサ・加速度センサを装着することで、容易に日常的な計測が可能である。JINS MEME は、慣性センサ・加速度センサを備えており、頭部運動の計測も可能である。本研究では、頭部・眼球運動の計測に JINS MEME を用いている。

2.2.2 視線移動の特徴による性格推定

視線移動の特徴による性格推定手法のうち、眼球運動を用いたものには、いくつかの既存研究がある [3] [4]。しかし、これらの既存研究では、視線移動の特徴から性格を推定するブラックボックスなモデルを用いており、視線移動の特徴と性格の関係性を具体的に可視化することはできない。

また、これらの既存研究における手法は、前述の 2 つの要件を満たしていない。まず、1) 眼球運動の特徴として、注視 (Fixation) や眼球運動 (サッカードなど) のイベントを手で検出し、それらの所要時間などを単純な特徴量として用いている。Berkovsky ら [3] は、実験室環境において、感情にはたらかける動画像を提示し、反応として現れる眼球運動の特徴量を人手で抽出している。このような特徴量を用いることで、Big Five を含んだ性格特性の強さを高い精度 (性格の高さを 3 クラスに分類し、最も精度のよい性格特性について 90.48%) で推定できることを報告している。一方で、Hoppe ら [4] は、大学の構内を歩き回り、売店で買い物をするという現実に近い課題を用いている。このような場合、同様に人手で抽出した眼球運動の特徴量から性格を推定すると、推定精度は高くない (性格の高さを 3 クラスに分類して、最も精度のよい性格特性について 48.6%)。すなわち、現実の作業を対象として、推定を精度よく行うためには、眼球運動の周期性や、頭部・眼球の協調運動など、人手では事前に想定しづらい、複雑な特徴を自動的にモデルに抽出させる必要があると推察される。本研究のように、性格と頭部・眼球運動の特徴の関係を分析することが目的の場合にも、このことは同様である。本研究では、特徴抽出を人手ではなく、モデルに自動で行わせることで、このような複雑な特徴も考慮することができる手法を提案している。

次に、2) これらの既存研究では、個人ごとの計測ノイズを分離していない。この点も、現実の作業を対象とした分析では課題となる。実際、計測ノイズが問題にならないほど統制された Berkovsky らの研究 [3] では、推定精度が高いが、計測ノイズが強い現実に近い Hoppe らの研究 [4] では、推定精度が低くなっている。Hoppe らの研究 [4] では、計測エラーが大きい約 2 割の実験参加者データを除外する必要があったことなどからも、個人ごとの計測ノイズを分離する必要性が推察される。本研究では、人手による特徴抽出不要で、行動の特徴を性格ラベル

(性格特性の強さ)、および個人差ラベル(個人ごとの計測ノイズ)の2つに分離して結びつけることができるモデルを提案している。

2.3 敵対的生成ネットワークと紐解き (Disentanglement)

2.3.1 敵対的生成ネットワーク

近年、深層学習を用いた手法により、人手による特徴抽出なしで、人の目に把握できる程度の低次元空間と、多次元データとを、相互に写像する研究が進展している。低次元から多次元への写像を行うアルゴリズムとして、潜在変数から多次元の観測データを生成するジェネレータと、生成された観測データと実データを識別するディスクリミネータからなる、敵対的生成ネットワーク (Generative Adversarial Networks: GAN) [8] が知られている。多次元から低次元への写像は、観測データから潜在変数を推定する経路を追加したディスクリミネータを用いた、相互情報量最大化敵対的生成ネットワーク (Information Maximizing GAN: InfoGAN) によって実現される。本研究では、InfoGAN によって、人手による特徴抽出なしで、多次元データである頭部・眼球運動データと、低次元空間である性格特性との関係を可視化する手法を提案している。

2.4 紐解き (Disentanglement)

潜在変数によって生成データの特徴を説明する際、特徴のばらつきを望んだ方法で潜在空間の次元に対応させることについては、紐解き (Disentanglement) と呼ばれる技術分野で研究が盛んである [9]。例えば、多次元の顔画像データを対象とした場合、顔の向きと表情を2次元の潜在空間の軸としてとることで、潜在空間上の座標を指定することで、顔の向きと表情を自由に変えながら多次元データである顔画像を生成することが可能となる。このように、顔の向きや表情といった意味を潜在変数に持たせるには、それらの意味を表すラベルを、潜在空間上の次元と一致させる教師あり学習が有効である。InfoGAN にこのような教師あり学習を導入する方法として、半教師あり相互情報量最大化敵対的生成ネットワーク (ss-InfoGAN) が提案されている [10]。本研究では、ss-InfoGAN によって、性格特性と個人差を別の潜在空間の次元として対応させることで、センサデータの特徴を、性格ラベル (性格特性の強さ)、および個人差ラベル (個人ごとの計測ノイズ) の2つに分離して結びつけることができる手法を提案している。

3 提案手法

提案手法では、ss-InfoGAN を用いて、JINS MEME¹によって取得された頭部・眼球運動のセンサデータと、Big Five 性格特性ラベル、個人差ラベル (実験参加者番号: ID)、経過時刻ラベル (Time) の対応関係をモデル化する。この手法により、センサデータと性格特性の対応関係が得られ、センサデータから各性格特性の強さを推定できるだけでなく、各性格特性の強さを変えながらセンサデータを生成可能となる。これにより、

性格特性の強さと関連して変化しているセンサデータの特徴を可視化・分析することが可能である。また、提案手法は、1) 人手による特徴抽出が不要であり、2) 性格特性とは別次元に個人差 (ID) を対応させることで、性格特性と個人ごとの計測ノイズを分離している。これらのことから、現実の作業を対象とする場合に必要となる2要件を満たしている。なお、経過時刻 (Time) を潜在変数に加えているのは、実験に参加した各個人の、課題を行っている各時刻について、ある性格特性の強さを変化させた場合のデータを生成することを可能とし、生成データに十分なバリエーションを持たせるためである。

以下では、提案手法における頭部・眼球運動データの前処理手法を説明し、その後、モデルの構築手法について説明する。モデルの構築手法についての説明では、準備として提案手法に用いている GAN, InfoGAN と、ss-InfoGAN について説明した後、提案手法の説明を行う。

3.1 提案手法のデータ前処理

3.1.1 頭部・眼球運動データ

頭部・眼球運動のセンサデータ (50Hz) を、入力データへ変換する方法を述べる。まず、センサデータを、計測ノイズを低減するため、時刻方向に長さ8ポイント (0.16秒) のハンニング窓を用いた移動平均によって平滑化する。次に、センサデータを128ポイント (2.56秒) 刻みで移動する512ポイント (10.24秒) の移動窓で分割し、センサデータの種別方向に9次元、時系列方向に512次元の入力データを得る。最後に、データの標準化を行う。提案手法で用いる GAN では、標準化により、入力データの値を $[-1, 1]$ の範囲に収めることが学習の安定化につながる [11]。しかし、頭部・眼球運動のセンサデータはばらつきが大きく、すべての時刻における値を範囲 $[-1, 1]$ に収めると、大部分の時刻において値が0周辺に集中してしまう。そこで、各実験参加者の、各課題について、各種別のセンサデータが、平均0、標準偏差1になるよう標準化する。

3.1.2 Big Five ラベル

Big Five の値は $[0, 28]$ の範囲をとるため、実験参加者全員の各特性値が平均0、標準偏差1となるよう標準化し、各入力データに対応した Big Five ラベルとして用意する。

3.1.3 実験参加者ラベル (ID)

本実験の参加者14名 (一部のデータに関しては13名) 分の配列を用意し、1つの要素が1、残りの要素が0となる one-hot ベクトルとし、各入力データに対応した実験参加者ラベル (ID) として用意する。

3.1.4 経過時刻ラベル (Time)

開始時点を-1、終了時点を1として、均等に値を増加させながら割り振り、各入力データに対応した実験中の経過時刻ラベル (Time) として用意する。

3.2 提案手法によるモデル構築への準備

3.2.1 GAN

GAN [8] では、ディスクリミネータは、自身に入力されるデータが、ジェネレータによって生成されたデータであるか、

¹ : <https://jins-meme.com/ja/>

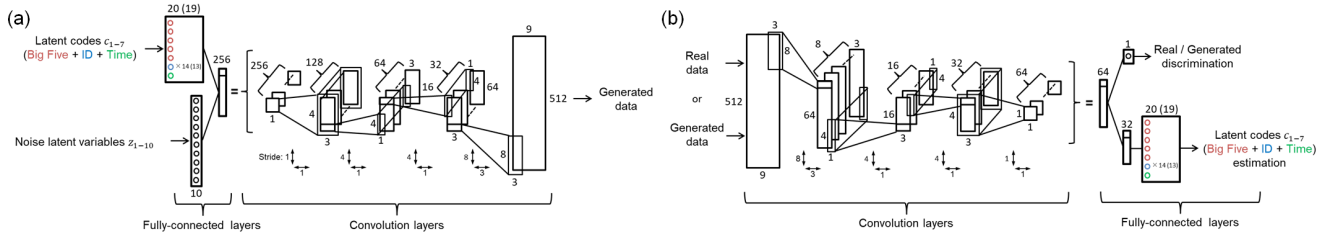


図 1 提案手法のアーキテクチャ (a) ジェネレータ (b) ディスクリミネータ

実データであるかを識別するように学習していく。一方で、ジェネレータは、自身が生成したデータが、ディスクリミネータによって識別されないように学習していく。このようなジェネレータとディスクリミネータの敵対的学習によって、ジェネレータにおいて、低次元の潜在変数から多次元データが生成される写像が得られる。

3.2.2 InfoGAN

InfoGAN [12] は、GAN のディスクリミネータに、潜在変数の推定を行う経路を追加したものである。InfoGAN では、ジェネレータは生成過程の元となる潜在変数 c 、ノイズ潜在変数 z を入力として、実データに類似したデータの生成を行う。ディスクリミネータでは、実データと生成データの識別に加えて、生成データの潜在変数 c の推定を行う。これにより、多次元データから低次元の潜在変数への写像が得られる。

3.2.3 ss-InfoGAN

ss-InfoGAN [10] とは、ジェネレータが生成したデータをディスクリミネータに入力する際には、既存の ss-InfoGAN と同様に潜在変数 c の推定を行うが、実データをディスクリミネータに入力する際には、実データに対応したラベルの推定を行うようにしたものである。これにより、潜在変数 c は、ラベルに紐づいた特徴を説明するようになる。

3.3 提案手法によるモデル構築

3.3.1 アーキテクチャ

提案手法のアーキテクチャを図 1 に示した。アーキテクチャは先行研究 [13] を参照して設計されている。ジェネレータは、1 つの全結合層と 4 つの逆畳み込み層で構成されている。潜在変数の次元数 20 あるいは 19 (内訳は Big Five 性格特性が 5 次元、ID が 14 あるいは 13 次元、経過時刻が 1 次元) と、ノイズ潜在変数の次元数 10 を加えた長さの配列からなる入力データは、多次元 (9×512) の生成センサデータに変換される。すべての中間層では、Relu function で活性化が行われる。出力層を除き、Batch Normalization を用いる。

ディスクリミネータは、全ての経路に共通した、4 つの畳み込み層の後に、経路ごとに異なる 1-2 層の全結合層で構成されている。センサ種別方向に 9 次元、時系列方向に 512 次元の入力データは、生成データと実データの識別を行う 1 次元の出力、および潜在変数の推定を行う 20 (19) 次元の出力に変換される。すべての中間層では、入力値域が 0 以下の場合 0.2 が乗算される Leaky Relu function で活性化が行われる。出力層を除き、全ての層で Spectral Normalization [14] を用いる。畳み込

み・逆畳み込み層においては、データサイズを合わせるための Reflection Padding を用いている。

3.3.2 学習

ランダムな順序でデータを用いる、サイズ 256 のミニバッチ学習を 100 エポック行う。それぞれのエポックにおいて、まずディスクリミネータの最適化を行い、次にジェネレータの最適化を行う。

ディスクリミネータの最適化過程を述べる。まず、ジェネレータを通してデータの生成を行う。ジェネレータの入力である潜在変数 c の次元数は 20 (19)、ノイズ潜在変数 z の次元数は 10 である。潜在変数 c のうち、Big Five に相当する c_{1-5} は平均値 0、標準偏差 1 の正規分布からサンプリングされる。実験参加者 ID に相当する c_6 は、14 (13) のユニットのうち 1 つが 1、残りが 0 となる one-hot ベクトルとし、どのユニットが 1 をとるかは $1/14$ ($1/13$) の確率でランダムにサンプリングされる。経過時刻 (Time) に相当する c_7 は $[-1, 1]$ の範囲を取る一様分布からサンプリングされる。ノイズ潜在変数 z は $[-1, 1]$ の範囲を取る一様分布からサンプリングされる。生成されたデータをディスクリミネータに入力し、「生成データ」「実データ」の判別を行う経路において、出力値をシグモイド関数に通した後に、交差エントロピー関数を用いて、「生成データ」として推定できるように誤差の評価・逆伝播を行う。次に、実データをディスクリミネータに入力し、「生成データ」「実データ」の判別を行う経路において、出力値をシグモイド関数に通した後に、交差エントロピー関数を用いて、「実データ」として推定できるように誤差の評価・逆伝播を行う。この際には、敵対的学習においてディスクリミネータがジェネレータに対して優勢になりすぎることを防ぐため、5%の確率で「実データ」「生成データ」のラベルを逆転させる [15]。一方、潜在変数 c を推定する経路では、潜在変数 c_{1-5} を対象として、入力データに対応した Big Five 特性の強さを推定するように、潜在変数 c_7 を対象として、入力データに対応した経過時刻 (Time) を推定するように、正規分布の負の対数尤度関数を用いて、誤差の評価・逆伝播を行う。潜在変数 c_6 を対象として、実験参加者 (ID) を推定するように、交差エントロピー関数を用いて、誤差の評価・逆伝播を行う。ディスクリミネータの最適化には、Adam ($\alpha = 0.002, \beta_1 = 0.1, \beta_2 = 0.5$) を用いる。

次に、ジェネレータの最適化過程を述べる。まず、ジェネレータを通してデータの生成を行う。生成過程については、ディスクリミネータの最適化の場合と同様である。生成されたデータをディスクリミネータに入力し、「生成データ」「実データ」の判

別を行う経路において、出力値をシグモイド関数に通した後に、交差エントロピー関数を用いて、「実データ」として推定されるように誤差の評価・逆伝播を行う。一方、潜在変数 c を推定する経路では、ジェネレータが生成したデータの元である潜在変数 c の値を推定できるように誤差の評価・逆伝播を行う。評価関数はディスクリミネータの最適化で述べた方法と同様である。学習は、ジェネレータに加えて、ディスクリミネータのうち潜在変数 c の推定にのみ関係している全結合層を対象に行う。ジェネレータの最適化には、Adam ($\alpha = 0.0002, \beta_1 = 0.1, \beta_2 = 0.5$) を用いる。これまでに述べた提案手法のパラメータは、予備実験によって選定されている。

4 評価実験

4.1 実験の方法

提案手法を評価するため、オフィス内を再現した課題中のセンサデータを実験参加者 14 名について取得した。具体的には、個人で行う課題として、会計書類を PC 上のファイルへ転記する課題、他者とのインタラクションが生じる課題として、実験進行者との会話を約 10 分間行うよう教示し、センサデータを計測した。加えて、休憩時の自由行動 10 分間のセンサデータを実験参加者 13 名について取得した（1 名分の減少は実験進行上の都合による）。センサデータの計測には、メガネ型ウェアラブルデバイス JINS MEME ESR² を用いた。頭部加速度 3 軸、頭部角速度 3 軸、水平・垂直方向の眼電位差、および左側電極の眼電位を 50Hz で計測した。

さらに、Big Five を測定する TDPI を用いて、実験参加者に自身の性格を評定してもらった。TDPI では、5 つの性格特性を計測するためにそれぞれ 4 つの質問が用いられており、計 20 の質問に自身が当てはまる程度を 7 段階で評定することで Big Five 特性の強さが取得された。

4.2 評価の方法

評価では、まず、学習が適切に進行していること、次に、モデルが頭部・眼球運動データから未知の性格特性を推定可能であることを検証した。最後に、モデルが推定可能であった性格特性の強さごとの特徴の可視化・分析を実施した。

4.2.1 学習の進行

提案手法が用いている ss-InfoGAN の学習では、大きく分けて 3 つの学習プロセス、すなわち、GAN におけるジェネレータとディスクリミネータの敵対的学習、InfoGAN における潜在変数 c の推定の学習、ss-InfoGAN における実データに対応したラベル推定の学習、が成り立っている必要がある。

まず、GAN の学習が成り立っていることを検証するため、学習過程におけるディスクリミネータが、ジェネレータが生成したデータに対して、一定の確率判定（出力層の値、Probability）で「実データ」と推定しており、かつ、実データに対しても、一定の確率判定で「実データ」と推定できていることを確認した。

次に、InfoGAN の学習が成り立っていることを検証するた

め、学習過程においてジェネレータに入力された潜在変数 c_{1-7} と、その潜在変数 c_{1-7} から生成されたデータをディスクリミネータに入力した際の出力である潜在変数 c_{1-7} の誤差が、学習の進行につれて減少したことを確認した。連続値である c_{1-5}, c_7 については最小二乗誤差（Mean Squared Error）で評価し、離散値である c_6 については正答率（Accuracy）で評価した。

最後に、ss-InfoGAN の学習が成り立っていることを検証するため、頭部・眼球運動の実データをディスクリミネータに入力した際の出力である潜在変数 c_{1-7} と、それらに対応した Big Five の値、実験参加者 ID、経過時刻（Time）の誤差が学習の進行につれて減少したことを確認した。評価方法は InfoGAN についての学習の評価と同様であった。

4.2.2 性格特性の推定

モデルが性格特性を推定できことを検証するために、転記・会話・自由行動において、バリデーションによる評価を行った。具体的には、転記・会話課題においてはモデルが頭部・眼球運動から性格特性を推定できるが、自由行動においては推定できないことを確認する評価を行った。

このような評価を行った理由を述べる。まず、転記・会話課題では、全ての実験参加者が同じ行動をとりつつも、その行動の特徴が性格特性ごとに異なって現れるはずである。そこで、転記・校閲課題で性格特性を推定できれば、モデルは性格ごとに異なる頭部・眼球運動の特徴を抽出したと考えられる。

一方で、自由行動では、実験参加者がとる行動が統制されていない。このような自由行動では、むしろ、バリデーションにおいて性格特性が推定できないことが重要である。提案手法は、個人ごとの計測ノイズを性格特性との関係から分離することができ、数人に共通して生じた計測ノイズを分離することはできない。例えば、性格特性の似た実験参加者は、メガネのかけ方が似通っており、眼電位に含まれるノイズの傾向が類似したなどの可能性がある。このような場合、提案手法は共通したノイズの特性を推定に用いてしまう可能性がある。この可能性を除外するために、自由行動におけるバリデーションを行った。上記のようなノイズは、行動が統制されていない自由行動中にも現れるはずである。すなわち、自由行動中にも性格特性が推定できた場合、モデルは数人に共通したノイズの差を根拠とした推定を行った可能性がある。その反対に、自由行動では性格特性が推定できないことを確認できれば、モデルは性格ごとに異なる頭部・眼球運動の特徴を抽出していたと考えられる。

具体的なバリデーション方法としては、leave-one-person-out cross validation（LOOCV）を用いた。LOOCV とは、バリデーション対象となる 1 人のデータを除いて学習を行ったモデルが、除かれた 1 人分のデータに対する推定が可能であるかを全員のデータを除外するまで繰り返して検証する方法である。Big Five 性格特性の値は、絶対値ではなく、他者と比較した相対的な値を分析することが重要であるため [16]、性格特性の推定値と実測値の絶対値の差ではなく、それら 2 者間の相関係数を検証した。相関係数が有意に正の値をとった特性について、モデルが推定可能であったとみなした。性格特性の推定値には、各個人の 10 分間の入力データから推定された性格特性の平均

2: <https://jins-meme.com/ja/researchers/specifications/>

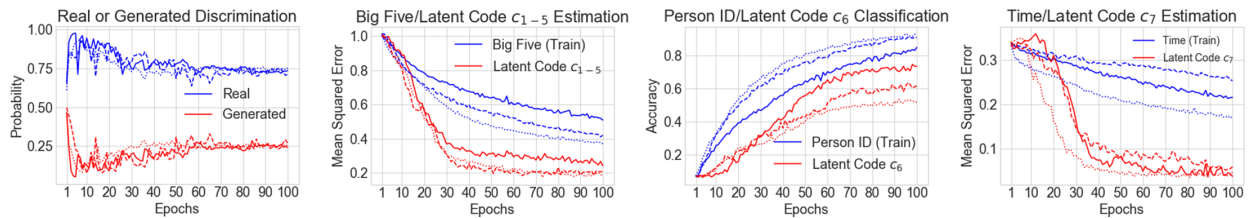


図2 学習における評価指標の変化（実線：転記，破線：会話，点線：自由行動）

値を用いた。

なお、LOOCVでは、バリデーション対象者1人のBig Fiveラベルは欠落させたが、頭部・眼球運動のセンサデータ、実験参加者ID、経過時刻（Time）は欠落させず利用した。これは、頭部・眼球運動のセンサデータから未知の作業者の性格を推定する場面を想定した時、推定対象者の頭部・眼球運動のセンサデータ、作業者ID、経過時刻（Time）は学習時から利用するためである。

4.2.3 性格ごとの特徴の可視化

性格ごとの特徴の可視化では、LOOCVの結果、モデルが推定可能であることが確認された性格特性を対象として、性格ごとの頭部・眼球運動の特徴の分析を行った。全ての実験参加者のデータを学習したモデルを用いて、性格特性の強さを変えながら頭部・眼球運動のセンサデータを生成し、特徴がどのように変化していったかを検証した。生成にあたっては、検証対象である性格特性に対応した潜在変数 c_{1-5} の値を-2から2まで0.4刻みで変化させながら、それ以外の潜在変数 c_i 、およびノイズ潜在変数 z については、学習時と同様のサンプリングを行って、100データずつ、計1100データを生成した。

本研究では、まずは簡易な分析までを行うこととし、評価しやすい水平方向の頭部・眼球運動として、生成データのうち、水平方向の頭部角速度センサ（Gyrometer Z）の値、および水平方向の眼電位差分（EOG H）を対象とした。生成データを視覚的に検証する予備調査から、分析では、時系列データの周期性を分析することが有効であると推察された。そこで、本分析では自己相関係数を指標として用いた。自己相関係数とは、時系列データ上のある時点の値を基準として、別の比較対象となる時点の値との相関を調べたものである。この相関を全ての時点について算出し、基準となる時点と比較対象となる時点の間隔（lag）がある特定の値の時に高くなるような場合、lagの大きさの周期性が認められると判断することができる。そこで、本評価では、上記2種の生成センサデータを対象として、lag 100（2秒）までの自己相関係数のグラフを描画し、性格特性の強さによってどのように周期性が変化したかを分析した。なお、自己相関係数の算出にあたっては、生成データを絶対値に変換する前処理を行っている。水平方向の頭部角速度センサ、水平方向の眼電位差分の値は、左右方向に運動が生じた際に一過性の増大あるいは減少を示す（図5上段、黒矢印部分）。そのため、この時系列データをそのまま用いて自己相関係数を算出した場合、一貫した方向への運動が生じた周期が算出されることになるが、どちらの方向に運動が生じるかには確率的なばらつ

きが大きく、周期性の算出が困難であると予想される。そこで、センサデータは平均0に標準化されていることに着目し、絶対値をとることで、左右どちらかの運動が生じた際に一貫して高い値をとる時系列データに変換した。

4.3 結果と考察

4.3.1 学習の進行

学習過程における評価指標のエポック毎の変化を図2に示した。紙面の都合から、全データを用いた学習のみを図示しているが、バリデーション時の学習でもほぼ同じグラフが得られた。評価指標の変化は、転記・会話・自由行動の全てにおいて、共通した傾向であった。GANの学習については、最終エポックのディスクリミネータの実データへの確率判定が70–80%、生成データへの確率判定が20–30%と、ディスクリミネータが優勢に偏ってはいるものの、後半のエポックにおいて両者が拮抗して安定的な状態となっていることから、敵対的な学習が成立していることが示唆された。

InfoGANの学習については、Big Five、Timeに対応する潜在変数 c_{1-5} 、 c_7 について、ほぼ誤差が消失するまで減少していた。IDに対応する潜在変数 c_6 については、若干、誤差の減少が鈍かった。ただし、14（13）カテゴリの分類問題であることを考慮すると、チャンスレベル（7–8%）を大きく上回った正答率が得られていることから（50–60%）、学習が進行したことが示唆された。

ss-InfoGANの学習については、Big Fiveに対応する潜在変数 c_{1-5} について、若干、誤差の減少が鈍かった。この誤差については、のちのバリデーションにおいて評価がなされるため、現段階では問題なく学習が進行したと考えた。IDに対応する潜在変数 c_6 については、ほぼ誤差が消失するまで減少していた。Timeに対応する潜在変数 c_7 については、誤差の減少がわずかであった。しかし、潜在変数 c_7 に与えられた役割は生成データのバリエーションを増やすことであるため、学習精度は大きく問われない。以上のことから、おおむね学習は適切に進行したと推察された。

4.3.2 性格特性の推定

LOOCVの結果、得られた各エポックの相関係数のグラフを図3に示した。転記・会話課題においては、学習が進行した後半のエポックにおいて、開放性の実測値と予測値の間で $r = 0.6$ 程度の強い正の相関がみられており、頭部・眼球運動は開放性と関連性を持つことが示唆された。また、最終エポックにおけるモデルの開放性の予測値と実測値の間の散布図を図4に示した。開放性についてのみ、相関係数は正に有意であった

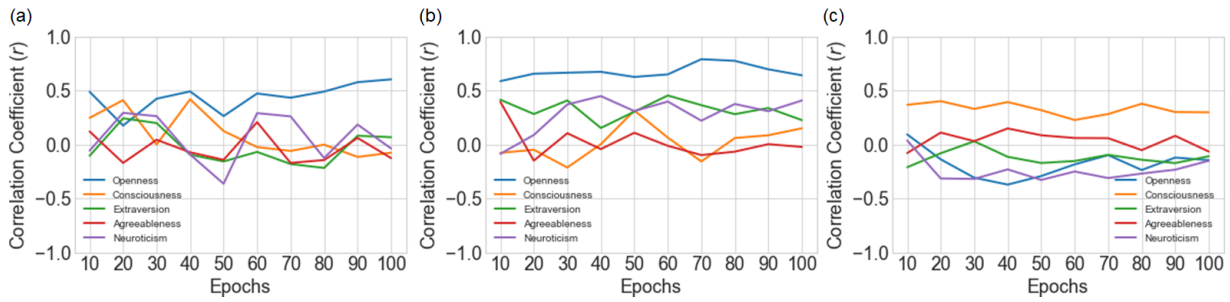


図3 各学習エポックにおける LOOCV の結果 (a) 転記 (b) 会話 (c) 自由行動

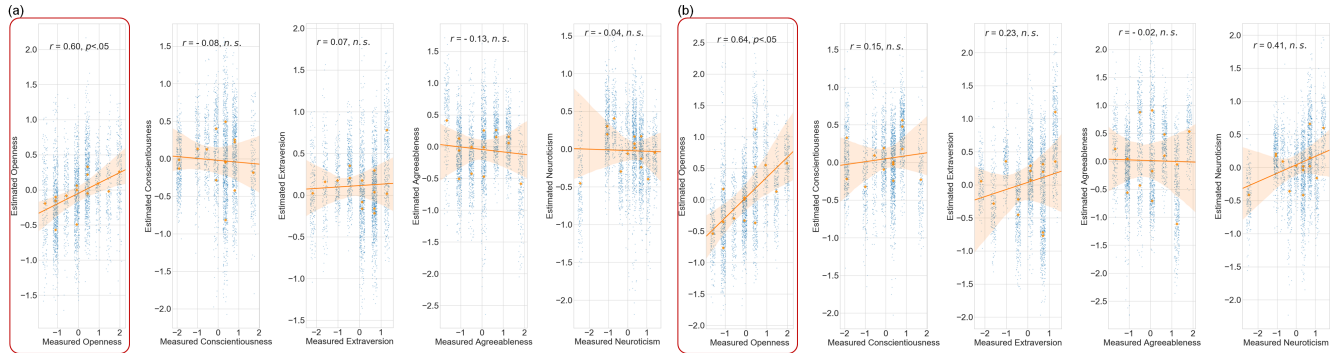


図4 最終エポックのモデルによる性格特性の推定値と実測値の散布図 (a) 転記 (b) 会話
青点は各データに対する推定値、橙点は個人ごとに平均化した推定値を表す³

($p < .05$, 図4)。

一方で、自由行動においては、開放性を含めた全ての性格特性について、予測値と実測値との間に相関は見られなかった。最終エポックにおける相関係数の値も、全ての性格特性について、有意なものはない ($n.s.$)。

以上のことから、本モデルは転記・会話中の作業者の開放性を推定可能であり、その推定は数人に共通した計測ノイズなどを根拠としたものではないことが示唆された。

4.3.3 性格ごとの特徴の可視化

最後に、転記・会話課題において、開放性の強さごとの頭部・眼球運動の特徴の分析を行った。まず、ランダムに選んだ水平方向の頭部角速度、水平方向の眼電位差の実データ、生成データを図5左に示した。生成データは、実データに比べノイズではあるものの、左右方向に運動が生じた際にのみ一過性の増大あるいは減少を示す（黒矢印部分）という、データの重要な性質を保存できていた。

開放性の強さを変化させながら生成したデータの、水平方向の頭部角速度、水平方向の眼電位差について、自己相関係数を算出した結果のグラフを図5右に示した。標準化済み開放性得点が低い場合 (-2.0 から -1.2) を再現した生成データのグラフは青線、中程度 (-0.8 から 0.8) の場合のグラフは緑線、高い場合 (1.2 から 2.0) のグラフは赤線である。グラフの極大地点が周期性を示すと考えられるため、それぞれのグラフに中で極大地点とみられる場所を青、緑、赤矢印で示した。図中に示すように、全体として開放性得点が高まるにつれて、極大地点（周

期性）がラグの大きな場所にずれていくような傾向が見て取れた。すなわち、転記・会話に共通して、開放性が高くなると、頭部・眼球運動の周期性が大きくなるという変化が見られた。

4.3.4 特徴可視化結果の考察

視線移動は、注視 (Fixation) と眼球運動 (Saccade) の繰り返しによって成り立っており、非常に大まかな区分をすれば、注視中には見ている情報の処理（視覚情報の意味を思考するなど）が行われ、眼球運動中には情報処理が行われないと考えられている [17]。実験室実験において、Big Five 性格特性と眼球運動の関係を調べた心理学分野における Rauthman らの研究 [1] では、眼球運動の単純な統計量と有意な関係を持っていた性格特性は開放性のみであると報告されている。この既存研究では、実験参加者に抽象的な画像を呈示した場合、眼球運動の統計量のうち、注視時間が増大すると報告されている。すなわち、開放性が高く、積極的に視覚情報の意味を考えながら課題を行った場合には、注視時間の眼球運動に対する時間割合が増大すると解釈された。このような時、眼球運動と頭部運動の発生周期は大きくなる。すなわち、本研究の結果は、Rauthman らの研究 [1] において示唆された性格と眼球運動の関係が、現実の作業においても生じていることを示唆したと考えられる。さらに、こうした関係が、眼球と協調している頭部運動においても現れること、また、視覚情報処理が中心的な課題ではない会話においても生じたことなどは、今後の性格と視覚情報処理の関係研究において、重要な知見となる可能性がある。

3：同一個人内での実測値のばらつきは、可視化のためランダムノイズを加えたことによる

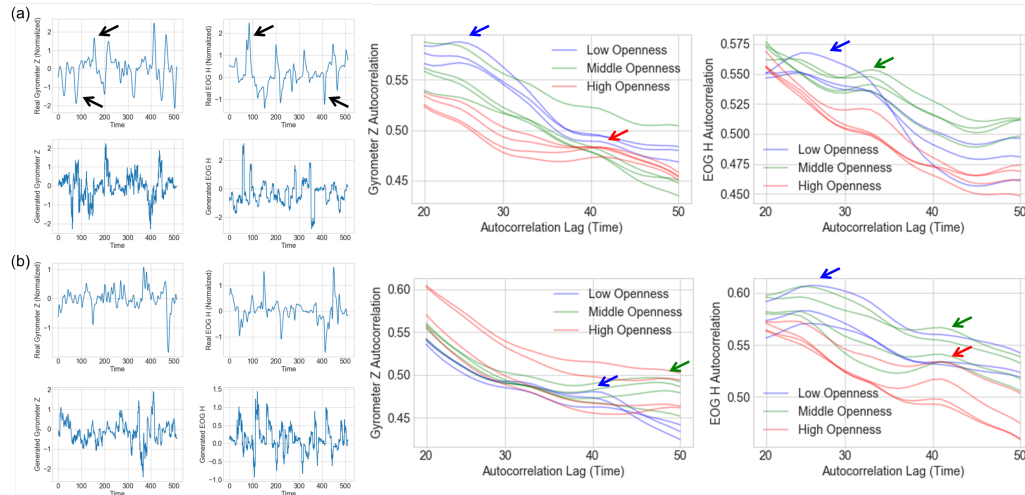


図 5 左：水平方向の頭部角速度，水平方向の眼電位差の実/生成データ (a) 転記 (b) 会話
右：水平方向の頭部角速度，水平方向の眼電位差の自己相関係数 (a) 転記 (b) 会話

5 結 論

本研究では，作業者がオフィス作業を行う際の頭部・眼球運動の特徴と作業者の性格の関係を，1) 人手による特徴抽出なしで，2) 個人ごとの計測ノイズを分離して，可視化できる手法を提案した．評価の結果，作業の内容が転記・会話課題として一貫して与えられた状況では，共通して Big Five 特性の開放性を推定することができた．また，作業の内容を特定しない自由行動中には，性格特性を推定できなかったことから，推定に計測ノイズなどが用いられていないことが示唆された．提案手法による可視化の結果，開放性が高くなると，頭部・眼球運動の周期性が大きくなることが示唆された．この特徴は，心理学分野における先行研究 [1] とも一致するものであった．

今後，生成データの分析を深めていくとともに，様々な作業場面に提案手法を適用することで，性格と特徴的な行動の関係を詳細に検証していく予定である．こうして得られた性格と行動の関係を活用し，作業者の現在割り振られている業務への好奇心の推定などにつなげ，働く人のモチベーションが向上する業務割り振りなどの実現を目指す予定である．

文 献

- [1] J. F. Rauthmann, C. T. Seubert, P. Sachse, and M. R. Furtner, "Eyes as windows to the soul: Gazing behavior is related to personality," *Journal of Research in Personality*, vol. 46, no. 2, pp. 147–156, 2012.
- [2] J. Yamashita, Y. Takimoto, H. Koya, H. Oishi, and T. Kumada, "Deep unsupervised activity visualization using head and eye movements," *IUI Companion*, pp. 41–42, 2020.
- [3] S. Berkovsky, R. Taib, I. Koprinska, I. Koprinska, E. Wang, Y. Zeng, J. Li, S. Kleitman, "Detecting Personality Traits Using Eye-Tracking Data", *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2019.
- [4] S. Hoppe, T. Loetscher, S. A. Morey, A. Bulling, "Eye Movements During Everyday Behavior Predict Personality Traits", *Front Hum Neurosci*. vol. 12, no. 105, 2018. Available: 10.3389/fnhum.2018.00105.
- [5] R. McCrae and P. Costa, "Validation of the five-factor model of personality across instruments and observers.", *Journal of Personality and Social Psychology*, vol. 52, no. 1, pp. 81–90, 1987. Available: 10.1037/0022-3514.52.1.81.
- [6] R. Iwai, D. Kawahara, T. Kumada, S. Kurohashi, "Development of a Japanese Personality Dictionary based on Psychological Methods", *Proceedings of the 12th Language Resources and Evaluation Conference*, pp. 3103–3108, 2020.
- [7] S. Ishimaru, K. Kunze, Y. Uema, K. Kise, M. Inami and K. Tanaka, "Smarter eyewear: using commercial EOG glasses for activity recognition," *UbiComp*, pp. 239–242, 2014.
- [8] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, and Y. Bengio, "Generative adversarial nets," *NIPS*, pp. 2672–2680, 2014.
- [9] H. Kim, and A. Mnih, "Disentangling by factorising," *arXiv preprint arXiv:1802.05983*, 2018.
- [10] A. Spurr, E. Aksan, and O. Hilliges, "Guiding infogan with semi-supervision," *ECML PKDD*, pp. 119–134, 2017.
- [11] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.
- [12] X. Chen, Y. Duan, R. Houthoofd, J. Schulman, I. Sutskever, and P. Abbeel, "Infogan: Interpretable representation learning by information maximizing generative adversarial nets," *NIPS*, pp. 2172–2180, 2016.
- [13] Z. Lin, K. K. Thekumparampil, G. Fanti, S. Oh, "InfoGAN-CR and ModelCentricity: Self-supervised Model Training and Selection for Disentangling GANs," *arXiv preprint arXiv:1906.06034*, 2020.
- [14] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida, "Spectral normalization for generative adversarial networks," *arXiv preprint arXiv:1802.05957*, 2018.
- [15] I. Goodfellow, "NIPS 2016 Tutorial: Generative Adversarial Networks," *arXiv preprint arXiv:1701.00160*, 2016.
- [16] 和田さゆり, "性格特性用語を用いた Big Five 尺度の作成," *心理学研究*, 67 巻, 1 号, pp. 61–67, 1996.
- [17] M. A. Just, P. Carpenter, "Eye fixations and cognitive processes," *Cognitive Psychology*, vol. 8, no. 4, pp. 441–480, 1976.