

作業現場におけるヒューマンエラーポテンシャルの 推定手法の検討

A study of human-error-potential estimation on a shop floor

大橋 洋輝[†]
Hiroki Ohashi

永吉 洋登[†]
Hiroto Nagayoshi

1 はじめに

少子高齢化が進行する日本や欧州では、高い技能を有する熟練労働者の減少が産業競争力の維持・向上への課題となって久しい。特に、人依存性の高い現場では、熟練労働者の減少によりヒューマンエラーが増加することが懸念され、生産性や品質の低下だけでなく、作業現場における重大事故やメガリコールが事業上の大きなリスクとして認識されるようになってきている。近年IoT関連技術の進展により、品質管理の4M (Man:人、Machine:機械、Material:材料、Method:方法) のデジタル化が進められているが、「人」の要素については、作業員ごとの差異はもちろんのこと、同じ作業員でも体調や精神状態などによって作業品質にばらつきが出ることが容易に予想され、その不確実性の大きさから十分にデジタル化が進んでいないと言えない。

本研究では、作業現場におけるヒューマンエラーの低減のため、まずその第一歩としてそれが発生するリスクの大きさ(以降これをヒューマンエラーポテンシャルと呼ぶ)を定量化することを目標としている。これをカメラや生体センサのセンシングデータから推定し、時々刻々変化する各作業員のヒューマンエラーポテンシャルを見える化することにより、短期的にはエラーポテンシャルが高くなってきている作業員に休憩を促したり、空調や照明をコントロールすることでヒューマンエラーを未然に防止することが、中長期的にはエラーポテンシャルと作業環境との関係性の統計分析を通じた作業環境の改善が期待できる。

これまでにも、集中力 [1] や疲労度 [2]、眠気 [3][4]、ストレス [5] など、人の内面状態をセンシングデータから推定する研究は数多くなされているが、これらの研究では人が着座して作業をしている状態、あるいは少なくとも大きく身体を動かしたりしない状況を暗に仮定している。一方、人の内面推定に用いられる生体指標の多くは身体動作によるノイズの影響を強く受けると考えられるため、これら既存研究の手法を本研究で想定するような作業現場にそのまま適用することは難しい。

そこで本研究では、身体動作を伴う状況下で正確にヒュー

マンエラーポテンシャルを推定する手法を開発することを主眼とし、その基礎検討を行う。生体指標の計算過程を確率的にモデル化し、生体指標の性質に関する知見を事前情報として組み込むとともに、カメラ映像や加速度センサから得られる「動き」に関連する情報を組み合わせることで、作業に伴う身体動作に起因するノイズが存在する状況下でも頑健な生体指標の推定を行うことのできる手法を提案する。研究室環境において10名の被験者により評価実験を行ったところ、70%超の精度でヒューマンエラーが発生しやすい状況を識別することができた。

2 関連研究

ヒューマンエラーの防止に関連する研究としては、認知科学や心理学、人間工学などの側面からその分類を試みるような研究が盛んになされている [6][7][8][9][10][11]。これらの研究では、ヒューマンエラーの分類結果や実際の事例に基づいた事前・事後の教育、あるいは比較的粒度の粗い情報に基づく工程改善などによるヒューマンエラーの防止策が議論・提案されている。作業現場に関するドメイン知識に基づいた定性的な分析やそれを受けた教育などは、現場の特性を踏まえた対策が可能であるため高い有効性が期待できる一方で、現場に慣れているからこそ見落としがちなリスクの存在や、顕在化こそしていないもののリスクとしては高い状態にあることなどは、何らかの客観的な指標をもって定量化すること無しに発見をすることは難しいと考えられる。ところが、既存研究の中では、時々刻々変化する作業員ごとのヒューマンエラーポテンシャルを客観的に定量化するような仕組みや、それに基づいた対策立案については十分な検討がなされていない。

一方でこれらの研究の中で、集中力や注意力、疲労度、ストレス、眠気など、人の内面的な要因がヒューマンエラーに強く関連していることが示唆されている。このような人の内的状態をセンシングによって推定する試みとしては、これまでも研究レベルのものから商用レベルのものまで、様々な研究がなされてきた。集中力を推定する手法としては、心電情報を用いるもの [12][13]、脳波情報を用いるもの [14]、皮膚電気活動情報を用いるもの [15]、指先温度を用いるもの [16]、加速度センサなどから得られる体動情報を用いるもの [17]、クッションの綿密度を用いるもの [18] や、それらを組み合わせたもの [19][20][21] などが存在する。Uema らの研究 [1] で

[†] 日立製作所 研究開発グループ
Research & Development Group, Hitachi, Ltd.

は、眼鏡型の眼電位センサから推定した瞬き情報を基に、人の集中状態を推定する手法が提案され、この眼鏡型のセンサは商用販売にも至っている。疲労度やストレス推定に関しても、同種のセンサが用いられることが多く、横山らの研究 [22] では心電センサから計算した心拍変動の時系列変化に着目して自動車運転時の疲労度推定が行われた。また松本ら [23] は同様の心拍変動解析を精神的ストレスの推定に用いた。居眠り運転防止のためにドライバの眠気を推定する研究も盛んになされており、車載カメラの画像を用いるもの [24] や、前述の心拍変動などと合わせた統合的な解析を行うもの [25] などが提案されている。

このように、これまでもヒューマンエラーに関連の強い集中力や疲労度など、人の内面を推定するような研究は様々な形でなされてきているが、これらの研究では暗に対象となる人が着座状態で作業をしていること、あるいは少なくとも大きく身体を動かしたりしない状況を想定しており、工場作業のように身体動作を伴うような作業状態をターゲットとしていない。身体動作を伴う状況では、心電センサや脳波センサなど各種生体センサに身体動作に伴うノイズが乗ることが予測され、既存手法をそのまま適用することは難しいと考えられる。間谷らの研究 [26] では、単調作業における作業者の体の動きに着目して、加速度センサによって繰り返し作業のリズムを推定し、周期的なリズムからの逸脱を検出することでヒューマンエラーの予兆を発見する手法が提案されたが、手法の評価のために行われた実験はタブレット PC 上でタッチペンで円をなぞるといったものであり、コンセプトとしては工場作業などを想定しているものの実際に行われた実験は着座作業であり、やはり身体動作を伴う作業への適用可能性についての検証はなされていない。

本研究では、工場や保守現場など、身体動作を伴う作業を行う場面でも、センシングデータからヒューマンエラーポテンシャルを正確に推定する手法を確立することをめざす。

3 ヒューマンエラーポテンシャルの定量化の方針

3.1 ヒューマンエラーの主な原因から考えるモデル化のアプローチ

ヒューマンエラーの分類としては、情報伝達のミス、感覚・認知のミス、判断・決定のミス、操作・動作のミスなど様々なものが挙げられるが [8]、このようなエラーのポテンシャルを推定するためには、作業工程や手順などある程度対象とする作業の内容に踏み込んだ解析を行う必要が生じると考えられる。そのようなモデル化を行うことで、対象とする作業についてはより精緻なモデル化ができると期待できる一方で、他の作業現場への汎用性は低下してしまう。そこで本研究では、様々な種別の作業に共通して存在し得る、より根本的な原因にフォーカスする。山本らの研究 [11] (特に表 3 を参照) によると、救急業務に関連した事故事例及びヒヤリハット事例において、その背後要因となっていたものは、主に「時間切迫に起因する焦り」「複数の事象に対する同時作業 (これを以後マルチタスクと呼ぶ)」「危険性の認知不足」の 3 つであった。

そこで本研究では、このようなヒューマンエラーが起りやすい状況を、カメラやウェアラブルセンサなどのセンシングデータから検知するモデルを構築することをめざす。より具体的には、センシングデータから、作業員が通常の状態で行っているのか、このようなヒューマンエラーが起りやすい状態で作業を行っているのかを識別するモデルを構築する。ただし、上記 3 つの主たる背後要因のうち、「危険性の認知不足」に関しては実験的に状況を再現することが難しいため、本研究では対象外とする。即ち、本研究では、ヒューマンエラーポテンシャルの推定問題を、「時間切迫状態」「マルチタスク状態」に「通常状態」を加えた、3 クラスの状態の識別問題としてモデル化する。

3.2 タスク設定

特定の作業に特化することなく、可能な限り汎用性の高いモデル構築を行うため、様々な作業現場で行われると考えられる工程からなる模擬作業を題材に実験を行うことにした。具体的には、「ねじ締め」「配線」「組付け」などの作業が含まれるデスクトップ PC のストレージ (SSD) 交換作業を題材とした。作業手順としては、デスクトップ PC の側面のカバーのねじを外す、デスクトップ PC の側面のカバーを外す、SSD が乗っているトレイを取り出す、SSD につながっているケーブルを抜く、SSD とトレイを固定しているねじをドライバーを使って取り外す、SSD を交換する、ドライバーを使って SSD をトレイにねじ止めする、SSD にケーブルを接続する、トレイを既定の位置に収納する、デスクトップの側面カバーを取り付ける、側面カバーをねじ止めする、である (図 1 参照)。尚、以降この一連の手順のことを「試行」と呼ぶ。

「時間切迫状態」の再現のためには、作業時間に制限を設け、被験者が作業を始める前にそれを伝えてその時間内に作業を終えるように指示した。また、作業中 30 秒ごと、残り 30 秒以下になってからは 10 秒ごとに経過時間を知らせ、より切迫感を生むようにした。「マルチタスク状態」の再現のためには、作業中に 2 桁の足し算・引き算の暗算問題を読み上げて出題し、これに声で回答しながら作業を行ってもらった。回答数や回答スピードについては特に条件を設けず、可能な限り作業と並行しながら暗算問題に答えてもらった。回答が間違っても再出題やペナルティを課すことはせず、そのまま次の問題に回答してもらった。

1 試行辺りの所要時間は、条件や被験者によっても異なるが約 3 分程度であった。

3.3 センサの選定

センサとしては、作業者に負担をかけることなくセンシングができ、かつ既に作業現場に配備が進んでいることも多い固定カメラに加え、既存研究で内面状態推定に一定の有効性が認められている心電・脈波・脳波・眼電・視線・加速度などのデータが取得できるウェアラブルセンサを用いる。皮膚電位等も有効性を認める報告はあるが、予備実験にて試行したところ、手指に電極を貼り付ける必要があり、著しく作業性を害したため、用いないこととした。具体的なセンサとしては、Logicool HD Pro Webcam C920 (固定カメラ、三人称映

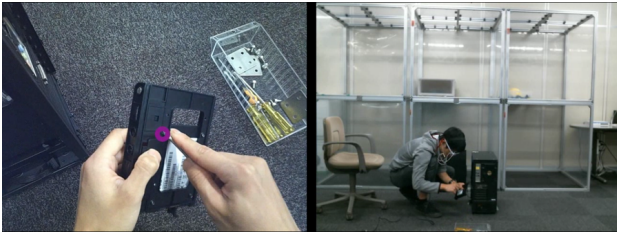


図1: SSD 交換タスクの様子。左: Eye tracking glass の映像 (中央の紫の円は推定された注視点位置、右: 固定カメラの映像)

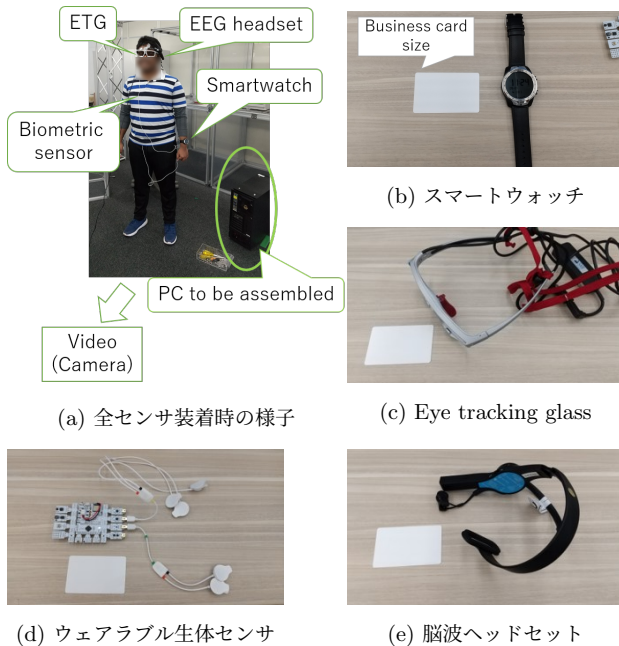


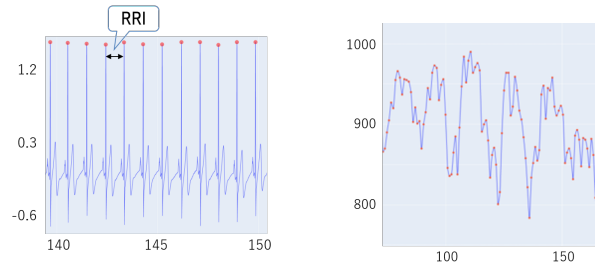
図2: 実験で使用したセンサ

像用)、SMI Eye Tracking Glass (一人称カメラ、一人称映像・視線情報用)、TicWatch Pro (スマートウォッチ、加速度データ・脈拍データ用)、biosignalsplux (ウェアラブル生体センサ、心電データ・眼電データ用)、MindWave Mobile 2 (脳波ヘッドセット、脳波データ用)を用いた。それぞれのセンサの外観と、その装着時の様子を図2に示す。

4 ヒューマンエラーポテンシャルの推定手法

4.1 特徴量抽出

本節では、ヒューマンエラーポテンシャルを推定するための特徴量抽出の方法について述べる。まず、4.1.1項で、全体的な方針について述べ、次に4.1.2項で特にノイズの影響下でも頑健に生体指標の算出を行うための確率モデルについて詳細に説明をする。4.1.3項では、身体動作のある状況下で、ヒューマンエラーポテンシャルの推定を行うために、生体指標と合わせて用いることを提案する身体動作に関する特徴量について説明をする。



(a) 安静時の ECG データの例。横軸: 時間 (秒), 縦軸: 電位 (mV)
(b) 安静時の RRI の例。横軸: 時間 (秒), 縦軸: 間隔 (ミリ秒)

図3: An example of ECG data and its RR interval (when a subject is staying calm)

4.1.1 各種センサデータからの特徴量抽出方針

心電位 (Electrocardiogram, ECG) データや眼電位 (Electrooculography, EOG) データ、脳電位 (Electroencephalography, EEG) データなどは、生データやその単なる時間平均などの統計量を解析に用いるのではなく、そこからピーク検出や周波数解析などを行って、心拍数や瞬き回数、特定周波数帯の脳波の強度などの特徴量に変換を行ってから解析することが一般的である。そこで本項では、3.3節で述べた各種センサから得られるデータの、基本的な特徴量抽出方針について説明をする。

心電位データについては、図3(a)に示すように心臓の鼓動リズムに対応した周期性の強い信号が観測される。このデータから R 波と呼ばれるピークに相当する位置 (図中赤丸) を求め、隣り合う R 波の間隔 (RR 間隔、RR interval; RRI) を算出する (図3(b)参照)。この RRI の試行ごとの平均値、RRI の時間変動を周波数解析することによって得られる LF (4Hz 以上 15Hz 未満の低周波成分)、HF (15Hz 以上の高周波成分)、LF/HF の値を特徴量として用いる。定性的には、LF は交感神経・副交感神経両方の活性度合いを、HF は副交感神経の活性度合いを表していると解釈される。尚、ノイズ下で頑健に RRI を算出する手法については、4.1.2 項で詳細に説明をする。

スマートウォッチから得られる脈拍情報についても同様の特徴量抽出を行う。実験で使用したスマートウォッチでは、心拍数 (60/RRI に相当) が 1 秒ごとに計測できるので、これを基に RRI を計算し、上記と同様にして RRI の平均値、LF、HF、LF/HF を求めて特徴量とする。

眼電位データについては、瞬きをしたときに鋭いピークが観測されるため、ピーク検出を行い試行ごとに瞬きの頻度 (回/分) を求めて特徴量とする。

Eye Tracking Glass (ETG) は、API により、画像内での注視点の位置、視線関連イベント (visual intake, saccade, blink) など、時刻ごとの視線に関する様々な情報を取得することができる。本研究では、注視点位置の x 軸方向 y 軸方向それぞれの標準偏差、連続するフレーム間における視線の移

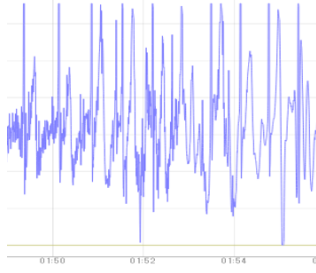


図4: SSD 交換タスク時の ECG データの例

動量の平均値及び標準偏差、視線が一定値以上動いた頻度、visual intake の割合、saccade の割合、を試行ごとに求め、これらを視線に関する特徴量とする。尚、イベント情報として得られる瞬き状態に関しては、事前に行った予備実験により、安定した精度が得られないことが分かったので、瞬き情報については前述の眼電位データからのみ取得することとした。

脳波データについては、Mindwave Mobile 2 の API により、 δ 波 (1-3 Hz)、 θ 波 (4-7 Hz)、低 α 波 (8-9 Hz)、高 α 波 (10-12 Hz)、低 β 波 (13-17 Hz)、高 β 波 (18-30 Hz)、低 γ 波 (31-40Hz)、中 γ 波 (41-50 Hz)、集中スコア、瞑想スコアが得られるので、これらの試行ごとの平均値及び標準偏差を特徴量とする。

スマートウォッチから得られる加速度データについては、4.1.3 項で述べる方法により、時刻ごとの身体動作を定量化した後、その時間平均を特徴量とする。また重力加速度の影響を軽減するため、一度各時刻の加速度の値から試行ごとの平均加速度の値を引いたものに対して、前述の定量化を施したものの時間平均も合わせて特徴量として使用する。

固定カメラから得られる映像データについては、4.1.3 項で述べる方法により、時刻ごとの身体動作を定量化した後、その時間平均を特徴量とする。

以上のようにして得られた値それぞれに対し、個人差によるバイアスを軽減するため、安静時の計測値からの増減値を特徴量として採用する。

4.1.2 ノイズ下での RRI の計算

心電データは、図 3(a) に示すように、安静時には非常に明確な周期信号が得られる一方、身体動作を伴う状況では、図 4 のように大きなノイズを伴うため、R 波を検出することは容易ではない。そこで、本研究では生体指標の計算過程を確率的にモデル化し、生体指標の性質に関する知見を事前情報として組み込む手法を提案する。

時刻 t における ECG の観測値を x_t とし、その時系列データの集合を $X_{t_s}^{t_e} = \{x_t\}_{t=t_s}^{t_e}$ とする。また、 n 番目の R 波の時刻を $t^{(n)}$ 、 n 番目の RRI の値を $y^{(n)}$ とおく (RRI の定義より、 $y^{(n)} = t^{(n+1)} - t^{(n)}$ の関係が成立する)。ECG の観測値列の長さを T とすると、ECG の観測値列 X_1^T と n 番目までの RRI が得られている時の $(n+1)$ 番目の RRI は次式によ

りモデル化することができる。

$$\begin{aligned} & p(y^{(n+1)} | X_1^T, y^{(n)}, \dots, y^{(1)}) \\ &= p(y^{(n+1)} | X_{t^{(n+1)}}^T, y^{(n)}, \dots, y^{(1)}) \quad (1) \\ &= \frac{1}{Z} p(X_{t^{(n+1)}}^T | y^{(n+1)}, y^{(n)}, \dots, y^{(1)}) p(y^{(n+1)}, y^{(n)}, \dots, y^{(1)}) \quad (2) \\ &= \frac{1}{Z'} p(X_{t^{(n+1)}}^T | y^{(n+1)}, y^{(n)}, \dots, y^{(1)}) p(y^{(n+1)} | y^{(n)}, \dots, y^{(1)}) \quad (3) \end{aligned}$$

ここで式 (1) においては、 n 番目の R 波が観測された後で $(n+1)$ 番目の R 波の時刻を求める際には、 n 番目の R 波の観測時刻以前の ECG の情報は寄与しないという仮定を用いた。また式 (2), (3) における Z, Z' は正規化のため定数である。

式 (3) 中の $p(X_{t^{(n+1)}}^T | y^{(n+1)}, y^{(n)}, \dots, y^{(1)})$ は、時刻 $\hat{t}^{(n+1)} = t^{(1)} + \sum_{i=1}^{n+1} y^{(i)}$ に R 波、即ち ECG のピークが観測されるとしたときに、ECG の観測列 $X_{\hat{t}^{(n+1)}}^T$ が得られる確率を意味し、この確率は時刻 $\hat{t}^{(n+1)}$ における ECG の値が大きいほど高いと考えられるので、次式のようにモデル化する。

$$p(X_{\hat{t}^{(n+1)}}^T | y^{(n+1)}, y^{(n)}, \dots, y^{(1)}) = \frac{1}{C_1} (x_{\hat{t}^{(n+1)}})^{\alpha}$$

ただし C_1 は正規化のための定数、 α はハイパーパラメータである。式 (3) 中の $p(y^{(n+1)} | y^{(n)}, \dots, y^{(1)})$ は n 番目までの RRI が得られている時に、 $(n+1)$ 番目の RRI が $y^{(n+1)}$ となる確率を表しており、本研究では以下のようにモデル化する。

$$\begin{aligned} & p(y^{(n+1)} | y^{(n)}, \dots, y^{(1)}) \\ &= \frac{1}{C_2} \left(\mathcal{N}(y^{(n)}, \sigma_1^2) + \beta \mathcal{N}(\mu^{(n)}, \sigma_2^2) + \gamma g(y^{(n+1)}) \right) \quad (4) \end{aligned}$$

ここで、 $\mathcal{N}(y^{(n)}, \sigma_1^2)$ 、 $\mathcal{N}(\mu^{(n)}, \sigma_2^2)$ はそれぞれ平均 $y^{(n)}$ 、分散 σ_1^2 及び、平均 $\mu^{(n)} = \frac{1}{n} \sum_{i=1}^n y_i$ 、分散 σ_2^2 の正規分布を指し、RRI が直前やこれまでの平均値と比べて急激に変化しないことをモデル化している。 C_2 は正規化定数、 β, γ はハイパーパラメータである。 $g(y^{(n+1)})$ は次式のように定義する。し、 n 番目までの RRI の観測値に関わらず当該 ECG 系列 X_1^T において RRI の値として $y^{(n+1)}$ が観測される確率を表す。

$$g(y^{(n+1)}) = \frac{1}{\sum_f h_{X_1^{t^{(n)}}}(f)} h_{X_1^{t^{(n)}}} \left(\frac{1}{y^{(n+1)}} \right)$$

ここで、 $h_{X_1^t}(f)$ は時系列データ X_1^t における周波数成分 f の信号の振幅を表し、 X_1^t に高速フーリエ変換 (FFT) を適用することによって求めることができる。

以上のようにして計算できる $p(y^{(n+1)} | X_1^T, y^{(n)}, \dots, y^{(1)})$ を用いて、 $(n+1)$ 番目の RRI $y^{(n+1)}$ は、以下のように求めることができる。

$$y^{(n+1)} = \arg \max_{\hat{y}^{(n+1)} \in Y^{(n+1)}} p(\hat{y}^{(n+1)} | X_1^T, y^{(n)}, \dots, y^{(1)})$$

ここで、 $Y^{(n+1)} = \{y^{(n+1)} | y_{min} < y^{(n+1)} < y_{max} \cap x_{t^{(n)}} + y^{(n+1)} \in M\}$ であり、 y_{min}, y_{max} は RRI として想定される最小及び最大の間隔、 M は ECG の極大値の集合である。定

性的には、 $Y^{(n+1)}$ は、 y_{min} から y_{max} の範囲の RRI のうち、 n 番目の R 波の時刻 $t^{(n)}$ から $y^{(n+1)}$ 経過後の ECG の値が極大値を取るような RRI の集合を意味する。

4.1.3 身体動作の定量化

スマートウォッチの加速度データおよび固定カメラの映像データを用いて、身体動作に起因する特徴量を定義し、これを用いることで、動き情報を加味したヒューマンエラーポテンシャルの推定を行う。

時刻 t における 3 軸加速度の測定値を $(a_{x,t}, a_{y,t}, a_{z,t})$ としたとき、加速度データから得られる身体動作特徴量 $m_t^{(acc)}$ を以下のように定義する。

$$m_t^{(acc)} = \sqrt{a_{x,t}^2 + a_{y,t}^2 + a_{z,t}^2}$$

また、固定カメラから身体動作特徴量を抽出するために、まず Pavlo ら [27] の手法を用いて、各時刻の主要関節の 3 次元位置の集合 $\{(l_t^{(1)}, \dots, l_t^{(j)})\}_{t=1}^T$ を求める。ここで、 $l_t^{(i)}$ は i 番目の関節の時刻 t における 3 次元座標 $(l_{x,t}^{(i)}, l_{y,t}^{(i)}, l_{z,t}^{(i)})$ であり、座標系は体の重心位置を基準とする相対座標となっている。また、 $J = 17$ である。次に、この 3 次元座標の時間差分を計算し、その絶対値を身体動作特徴量とする。即ち、固定カメラから得られる関節 j に関する身体動作特徴量 $m_t^{(video,j)}$ を以下のように定義する。

$$m_t^{(video,j)} = \sqrt{(l_{x,t}^{(j)} - l_{x,t-1}^{(j)})^2 + (l_{y,t}^{(j)} - l_{y,t-1}^{(j)})^2 + (l_{z,t}^{(j)} - l_{z,t-1}^{(j)})^2}$$

4.2 特徴量選択

4.1 項で述べた手順により 1 試行あたり合計 55 個の特徴量が得られる。十分な数の学習データが得られれば、全ての特徴量を用いてモデルを学習することが可能であると期待できるが、学習データ量が少ない場合、過学習を起こす可能性が高い。そこで、各特徴量を個別に分析し、タスク実行条件の種別の違い (通常条件、時間切迫条件、マルチタスク条件) によって一定の傾向が見られるもののみを採用することにした。

分析の結果、図 5 に示すようにの 10 種の特徴量に一定の傾向が見られたため、これらの特徴量として採用することにした。尚、一人の被験者については、視力の関係から ETG を装着しての実験ができなかったため、当該被験者については Visual intake や Saccade の特徴量が得られていない。

4.3 識別モデル

識別モデルとしてはニューラルネットワークを用いる。4.2 で述べた特徴量間には局所性や構造的な特徴は仮定できないため、CNN などではなく、全結合層のみを用いることとした。最適化ライブラリ Optuna [28] を用いて TPE アルゴリズム [29] により最適化を行い、ネットワークの隠れ層の数は 3、それぞれのノード数は 200, 50, 50 とした。初期の学習率は 0.03 とし、 t エポック目の学習ではそれを $1/t^{0.5}$ 倍して学習率を減衰させた。

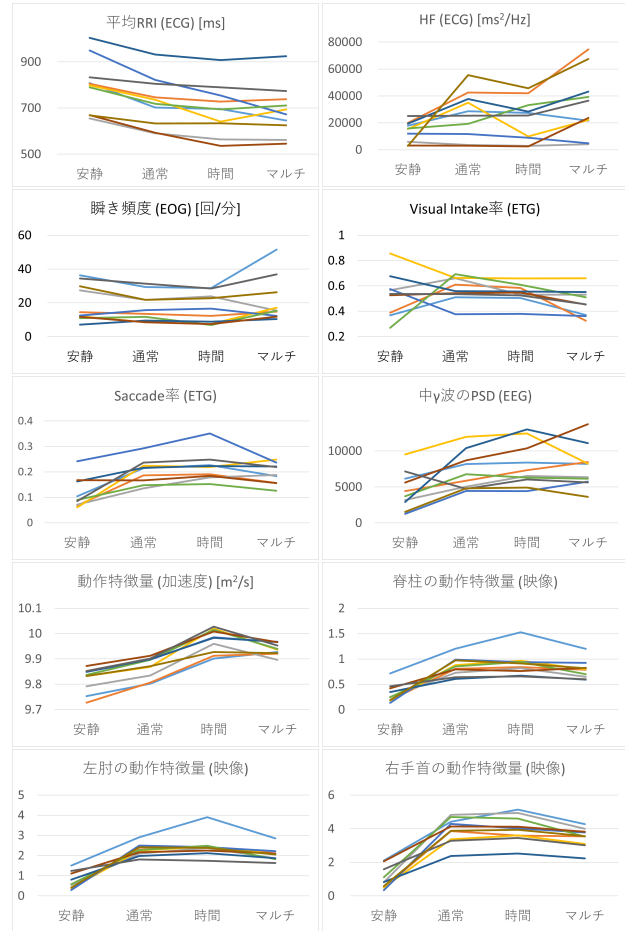


図 5: 選択した特徴量。色の違いは被験者の違いに対応。

5 評価

5.1 実験設定

提案手法の有効性の評価のため、10 名の被験者を募り、実験を行った。尚、被験者 10 名は実際の工場などで働く作業員ではなく研究員であり、実験は研究室環境にて行った。実験では、まず被験者に実験の主旨や内容について説明し、次に作業に慣れてもらうためにセンサをつけない状態で SSD 交換タスクの練習をしてもらった。十分にタスクに慣れたと被験者が感じた後、各種ウェアラブルセンサを取り付け、全てのセンサが正常に機能していることを確認した後、安静時状態の計測を行った。その後、センサをつけた状態で再度 SSD 交換タスクの練習をもらい、本番の計測を行った。本番の計測では、被験者ごとに通常条件、時間切迫条件、マルチタスク条件それぞれ 3 試行、合計 9 試行分の計測を行った。1 試行が終了するごとに休憩を挟み、その時同時に安静時の状態の計測も行った。尚、3 種の条件をどのような順番で行うかは、カウンタバランスを取るために被験者ごとに変更した。一例として、被験者 S1 は通常 → (安静 →) 時間切迫 → (安静 →) マルチタスク → (安静 →) 通常 → … → マルチタスク、といった順で、被験者 S2 は時間切迫 → (安静 →) マルチタスク

表 1: 評価の結果。通常：通常条件、時間：時間切迫条件、マルチ：マルチタスク条件

		推定値			
		通常	時間	マルチ	計
真値	通常	7.3	2.3	0.3	10
	時間	2.3	6.7	1.0	10
	マルチ	0.3	1.3	8.3	10
	計	10.0	10.3	9.7	30

表 2: 使用特徴量による精度比較。3クラス：通常、時間切迫、マルチタスク条件の分類、2クラス：通常条件とヒューマンエラーの生じやすい状況との二値分類。

#	特徴選択	動作特徴	3クラス	2クラス
1			56.5	71.1
2	✓		70.0	78.9
3		✓	61.1	65.6
4	✓	✓	74.4	82.2

→ (安静→) 通常→ (安静→) 時間切迫→...→通常、といった順で行った。各試行は通常約2~4分程度、実験全体で1被験者当たり約2時間程度の時間を要した。

5.2 評価方法

評価は各被験者の各条件3試行分を3つに分割し、3-foldのクロスバリデーションを行った。つまり、学習データには各被験者各条件の計測データが2試行分ずつ、合計で被験者当たり6試行分、総合計で60試行分が含まれ、評価用データには残りの30試行分が含まれている。評価指標はこの3-foldのクロスバリデーションにおける平均精度とした。

5.3 結果と考察

表1に提案手法による精度評価の結果を示す。総合的な精度は74.4% (=67/90)、通常条件とヒューマンエラーの生じやすい状況 (つまり時間切迫及びマルチタスク条件) との二値分類問題としては、82.2% (=74/90) の精度となった。

表2は使用特徴量による精度の比較結果である。#1及び#2は4.1.3で述べた身体動作特徴量を用いずに、生体センサから得られる特徴量のみを用いた場合の精度である。特に、#1は全ての生体特徴量を用いた結果、#2は4.2項で選出した10種の特徴量のうち、身体動作に関わるものを除いた場合の結果である。これらが、#3及び#4と比べて精度が低下していることから、身体動作特徴量を用いることの重要性が確認できる。これは、本研究のように動きを伴う作業中にヒューマンエラーポテンシャルのような人間の内面に関する指標を求める際には、身体動作を勘案することが重要であることを示唆している。また、特徴量選択を行わずに55個の特徴量全てを用いた#3の精度があまり高くないことから、特徴量選択を行わないと、過学習が生じ精度が低下していることが分かる。

6 結言

本研究では、工場や保守現場などにおける作業中のヒューマンエラー対策支援のために、カメラやウェアラブルセンサのデータからヒューマンエラーポテンシャルを定量化する手法の初期検討を行った。身体動作を伴う状況下でヒューマンエラーポテンシャルを推定するために、生体指標の計算過程を確率的にモデル化し、生体指標の性質に関する知見を事前情報として組み込むとともに、カメラ映像や加速度センサから得られる「動き」に関連する情報を組み合わせる手法を提案した。10名の被験者により評価実験を行ったところ、70%超の精度でヒューマンエラーが発生しやすい状況を識別することができた。

参考文献

- [1] Y. Uema and K. Inoue. JINS MEME Algorithm for Estimation and Tracking of Concentration of Users. In *Proceedings of UbiComp/ISWC*, 2017.
- [2] G. Sikander and S. Anwar. Driver Fatigue Detection Systems: A Review. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 20, No. 6, pp. 2339–2352, 2019.
- [3] M. Tsujikawa, Y. Onishi, Y. Kiuchi, T. Ogatsu, A. Nishino, and S. Hashimoto. Drowsiness estimation from low-frame-rate facial videos using eyelid variability features. In *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018.
- [4] M. Sun, M. Tsujikawa, Y. Onishi, X. Ma, A. Nishino, and S. Hashimoto. A neural-network-based investigation of eye-related movements for accurate drowsiness estimation. In *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018.
- [5] S. Kido, A. Hashizume, T. Baba, and T. Matsui. Development and evaluation of a smartphone application for self-estimation of daily mental stress level. *International Journal of Affective Engineering*, Vol. 15, No. 2, pp. 183–187, 2016.
- [6] A. Hale, J. Stoop, and J. Hommels. Human error models as predictors of accident scenarios for designers in road transport systems. *Ergonomics*, Vol. 33, No. 10–11, pp. 1377–1387, 1990.
- [7] 小松原明哲. 認知人間工学からのベテラン作業者のヒューマンエラーの防止. *安全工学*, Vol. 38, No. 6, pp. 352–358, 1999.
- [8] 佐谷克明. ヒューマンエラー防止へのアプローチ. *安全工学*, Vol. 38, No. 6, pp. 380–388, 1999.
- [9] A. C. Edmondson. Learning from mistakes is easier said than done: Group and organizational influences

- on the detection and correction of human error. *The Journal of Applied Behavioral Science*, Vol. 40, No. 1, pp. 66–90, 2004.
- [10] 篠原一光. 認知心理学から見たヒューマンエラー. *Medical Gases*, Vol. 17, No. 1, pp. 7–13, 2015.
- [11] 山本建太郎. 救急業務に関連した事故事例及びヒヤリハット事例における心理的要因の分析. 日本救急医学会関東地方会雑誌, Vol. 39, No. 3, pp. 405–408, 2019.
- [12] 成澤元, 青木美聡, 高橋敏治. チューニングによる計算課題への集中力促進効果. 日本心理学会第78回大会, 2014.
- [13] 駒澤真人, 板生研一, 菅谷みどり. 心拍変動を活用した集中状態の評価検討. 信学技報, Vol. 119, No. 44, pp. 73–76, 2019.
- [14] 稲田脩二, 早川孝之, 黄宏軒, 原田史子, 島川博光. 脳波によるヒューマンエラーの予兆となる注意力の低下の検知. 情報科学技術フォーラム講演論文集, Vol. 12, No. 4, pp. 627–630, 2013.
- [15] 志水雅俊, 棟方渚, 松原仁. 皮膚導電率反応による集中の持続とパフォーマンスの分析. 情報処理学会 第74回全国大会講演論文集, Vol. 2012, No. 1, pp. 145–146, 2012.
- [16] 新島有信, 藪内勉, 篠原章夫, 定方徹, 小林稔. 指先皮膚温度変化に基づく集中力推定方法に関する一考察. 電子情報通信学会技術研究報告: 信学技報, Vol. 112, No. 474, pp. 11–16, 2013.
- [17] 佐藤信夫, 辻聡美, 矢野和男. ビジネス顕微鏡を用いた個人作業時における集中状態判定特徴量の開発. 情報科学技術フォーラム講演論文集, Vol. 11, No. 4, pp. 65–68, 2012.
- [18] 鶴岡秀樹, 後藤隆太郎, 横田悠右, 成瀬康, 矢入郁子. クッション型デバイスによる着座姿勢と学習集中度の評価. 人工知能学会 全国大会論文集, 2017.
- [19] 内田昂, 磯山直也. 脈波及び体動情報を用いた集中状態推定精度の検証. マルチメディア, 分散協調とモバイルシンポジウム 2017 論文集, Vol. 2017, pp. 876–879, 2017.
- [20] 濱谷尚志, 内山彰, 東野輝夫. 種々のセンサを併用した集中度センシング法の検討. 研究報告高度交通システムとスマートコミュニティ (ITS), Vol. 2015, No. 10, pp. 1–6, 2015.
- [21] 高橋裕也, 駒澤真人, 菅谷みどり. 脳波と心拍変動を活用した集中状態の評価検討. 信学技報, Vol. 119, No. 225, pp. 19–21, 2019.
- [22] 横山清子, 高橋一誠. 心拍変動時系列による自動車運転時の主観的疲労感推定の基礎的検討. 電子情報通信学会論文誌 A, Vol. 96, No. 11, pp. 756–762, 2013.
- [23] 松本佳昭, 森信彰, 三田尻涼, 江鐘偉. 心拍揺らぎによる精神的ストレス評価法に関する研究-パターン認識による評価法の考察. ライフサポート, Vol. 24, No. 2, pp. 62–69, 2012.
- [24] 佐藤優伍, 野澤直樹, 森島繁生. 顔画像特徴と眠気の相関に基づくドライバーの眠気検出. 情報処理学会 第78回全国大会講演論文集, Vol. 2016, No. 1, pp. 331–332, 2016.
- [25] 今井章博, 小栗宏次. 居眠り運転時の特徴的な眠気推移傾向を組み込んだ眠気レベル推定手法の提案. 電子情報通信学会論文誌 D, Vol. 96, No. 4, pp. 1012–1019, 2013.
- [26] 問谷洋平, 梶原祐輔, 島川博光. 単調作業における作業者の体の動きに着目したヒューマンエラー予兆の発見. 情報科学技術フォーラム講演論文集, Vol. 13, No. 3, pp. 313–314, 2014.
- [27] D. Pavllo, C. Feichtenhofer, D. Grangier, and M. Auli. 3D human pose estimation in video with temporal convolutions and semi-supervised training. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.
- [28] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama. Optuna: A next-generation hyperparameter optimization framework. In *International conference on knowledge discovery & data mining*, 2019.
- [29] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl. Algorithms for hyper-parameter optimization. In *Neural Information Processing Systems (NIPS)*, 2011.