

スマートフォンのセンサ情報を用いた ユーザのストレス検知手法の検討

阿部 和樹[†] 田村 柁優紀[†] 中村 聡史^{††} 山中 祥太^{†††}

[†] 明治大学大学院先端数理科学研究科 〒164-8525 東京都中野区中野 4-21-1

^{††} 明治大学総合数理学部 〒164-8525 東京都中野区中野 4-21-1

^{†††} ヤフー株式会社 〒102-8282 東京都千代田区紀尾井町 1-3 東京ガーデンテラス紀尾井町紀尾井タワー

E-mail: [†] ai3ekzk@gmail.com, mogamusa31@gmail.com ^{††} satoshi@snakamura.org

^{†††} syamanak@yahoo-corp.jp

あらまし スマートフォンで利用するアプリや Web サービスには、使いづらい UI や画面読み込みによる遅延など、ユーザがストレスを感じる場面が多数存在する。これらの問題に対してサービスの開発者はユーザテストやフィードバックから問題箇所の特特定などを行うが、ストレスを感じる場面の検知は人手で行われるため、サービスにおける問題箇所を完全に特定することは難しい。そのため、ユーザがストレスを感じる場面を人手ではなく自動で検知することにより、開発者がサービスの問題点をより特定しやすくなると考えた。そこで本研究ではユーザがストレスを感じた際の身体の動きに着目し、それらの特徴をスマートフォンのセンサによって捉えることで、ユーザがストレスを感じたかどうかを判定する手法を検討する。具体的には、スマートフォンに搭載されている加速度センサ、角速度センサを利用し、それらのデータの特徴から機械学習によってユーザのストレス状態の判定を行った。判定実験の結果、69.1%の精度で判定できることがわかったが、より正確にストレス状態を判定するためには心拍など定量的にストレスを計測できる方法によってストレス状態を明確にする必要があるなどの課題が見つかった。

キーワード ストレス、スマートフォン、センシング、機械学習

1. はじめに

スマートフォンは日常の様々な場面で活用されており、多種多様なアプリ・Web サービスがスマートフォンによって利用されている。しかし、それらスマートフォンにおけるサービスの利用において、ユーザがストレスを感じる場面が多数存在する。例えば、画面の読み込みに時間がかかる場面、Web ページに頻出するバナー広告などの意図しない部分をタッチしてしまう場面、画面上のテキストのコピー範囲をうまく指定できないなどの場面がある。

このようなストレスを感じる場面を減らすためには、サービスの使いやすさ（ユーザビリティ）を向上させる必要がある。サービスの開発者は、これらのユーザビリティの向上のため、サービスをリリースする前のユーザテストや、実際に利用しているユーザからのフィードバックによってサービスの改善を行う。しかし、これらは人手によってサービスの問題点を探し出しているため、すべての問題点を把握するには膨大な労力が必要であるうえに、細かい問題点などは補足しきれない可能性が考えられる。

そのため、ユーザがストレスを感じる場面を自動的に検出することができれば、サービスにおけるユーザビリティの向上に繋がると考えられる。ここで、人のストレスを検出する研究としては、心拍数や発汗などの生理的反応からストレスを測定する研究[1]が存在

する。しかし、これらの手法は専用の装置を使用することや、測定のための環境が必要となるため、実際にサービスを利用する場面とは状況が異なると考えられる。また、測定するための装置の装着により、ユーザの体験が損なわれるなど、意図しないストレスの場面が発生することも考えられる。そのため、ストレスを検出する仕組みをサービスに手軽に組み込めることが望ましいと考えられる。

ここで、ユーザがスマートフォンのサービスに対してストレスを感じる場面において、ユーザが特徴的な行動を起こすことが考えられる。例えば、イライラすることによって乱暴な操作になってしまうことや、同じボタンを繰り返してタップすること、不必要な画面の操作を行うことなどの行動が考えられる。このような、平常時とは異なるスマートフォンの操作を自動で検出することができれば、ユーザがストレスを感じる場面を発見することに繋がると考えられる。

そこで我々は、これらの平常時とは異なる行動の検出について、スマートフォンの動きに関する情報に着目する。本研究では、スマートフォンの動きに関する情報として、スマートフォンに搭載されている加速度センサ、角速度センサから得られる情報を利用し、それらをもとにユーザがストレスを感じる場面を検出する方法を検討する。実際にユーザがサービスに対してストレスを感じる場面を作り、ストレスを感じる場面

におけるセンサ情報を収集する。また、それらの情報をもとに、機械学習によってユーザのストレス状態を推定する方法を試みる。

2. 関連研究

ユーザビリティの向上のため、サービスの問題点を発見するための様々な研究が存在する。例えば、ユーザビリティを評価する方法の1つとして、ユーザビリティに関するガイドラインや専門家の意見をもとに評価を行うヒューリスティック評価法[2]が存在する。また、坂井らの研究[3]では、ユーザの視線ログのデータから Web サービスのユーザビリティを評価する方法を提案している。これらに対して、コストや労力をかけずにサービスの問題点を検出するため、Paterno らの研究[4]ではサービスの利用におけるユーザの行動の記録から問題点を検出するツールの開発を行っている。また、Eardley らの研究[5]では、使いづらいユーザインタフェースにおけるスマートフォンの操作について、ユーザの手の使い方やスマートフォンのサイズによる操作方法の変化を調査している。また、Tajima らの研究[6]では、Web ページの入力フォームにおいて入力ミスを起こしやすいユーザインタフェースに対し、それらの入力ミスを自動でユーザに提示する方法を提案している。これらの研究のように、サービスの問題点を発見するための様々な方法が存在する。本研究では、このようなサービスの問題点を検出する方法の1つとして、ユーザがサービスにストレスを感じる場面を発見することを目的とする。

また、ユーザのストレスを検出するための様々な研究が存在する。ストレスは人の自律神経の変化が関係するため、それらを生理的反応から計測する方法として、心拍数・発汗・呼吸などによって測る方法が存在する[1]。また、特定の装置を使うことで、ストレスを計測する研究も存在する。小川らの研究[7]ではメガネ型デバイスである JINS MEME[8]を利用して、精神的負荷がかかる作業場面を眼電位のデータから推定する方法を提案している。また、Hernandez らの研究[9]では、PC を利用するユーザについて、センサを搭載したマウスとキーボードを用意し、それらのデバイスから得られる情報とストレスの関係を調査している。このように、ストレスを検出するための様々な方法が存在するが、ストレスを測るための専用の装置が必要であるなど、実際にサービスを利用する場面に適していないという問題がある。本研究では、スマートフォンのセンサ情報を利用し、サービスに容易に組み込むことができるストレス検出方法を検討する。

3. データセット構築

本研究では、スマートフォンの動きに関するセンサ情報をもとに、ユーザがストレスを感じている場面を検出することを目的とする。そこで、実際にスマートフォンの利用においてストレスを感じる環境を作り、それらを利用している間のセンサ情報を取得する。3章ではこれらのセンサ情報を集約したデータセット構築について説明し、4章ではデータからストレスを感じている場面を検出可能かについて判定実験を行う。

3.1. データ収集タスク

データセット構築のため、ストレスを感じる場面について実際のサービスで起こりうる状況を再現したデータ収集タスクを用意する。

今回はデータ収集タスクとして、漫画ビューアによって漫画を読むタスクを用意し、読んでいる途中にページ読み込みによる遅延が発生する環境を用意した。タスクでは、実験協力者にスマートフォンによって漫画を1冊読んでもらう。その際、ページの遷移時に強制的にページの読み込みを発生させ、普通より遅れてページが表示するようにした。遅延発生時のページ遷移の様子を図1に示す。このページの読み込みは、通信速度が十分でない状況において画像の読み込みが遅れてしまうという実際の状況を想定している。このようなサービスにおける情報の読み込み(ローディング)は、スマートフォンのスムーズな利用の妨げとなるため、ユーザがストレスを感じやすい場面であると考えられる。また、漫画のように続きが気になるコンテンツの閲覧時に、読み込みによって待たされることで、よりストレスを感じやすい環境を作ることができると考えられるため、データ収集タスクとして選定した。なお、今回のデータ収集タスクでは実際に通信速度を遅くするのではなく、プログラム側の操作によって強制的に遅延を発生させる。



©赤松健『ラブひな』(Manga109[10])

※タスクで実際に使用した漫画とは異なる

図1 遅延発生時のページ遷移

タスクで使用した漫画作品は1巻で完結となる作品を使用した。また、遅延が発生する状況として、連続する6ページにおける5回のページ遷移時にそれぞれ遅延を起こした。これは、漫画の先が気になる状況において立て続けに遅延することにより、ストレスが増幅するのではないかと考えたためである。遅延時間は最大10秒のランダムな時間で発生するように設定した。これらの連続するページにおける遅延発生を、1冊あたり最大10シーン発生させるため、5回×10シーンの最大50回の遅延発生のタイミングが存在する。

3.2. 構築内容

前節で述べたデータ収集タスクによって、ユーザがストレスを感じる場面におけるスマートフォンのセンサ情報を収集する。センサ情報としては、スマートフォンに搭載されている加速度センサ、角速度センサの情報に加え、それらのセンサ情報から計算される姿勢情報（スマートフォンの回転方向の情報）の3種類の情報を取得し、データとして記録した。ここで、加速度センサ、角速度センサではそれぞれX軸、Y軸、Z軸の3軸の情報があり、姿勢情報ではロール（X軸の回転方向）、ピッチ（Y軸の回転方向）の2軸の情報が存在する。そのため、取得するデータとして合計8種類の数値が存在する。スマートフォンにおける各データの扱いを図2に示す。データの記録は漫画の読書中に常に行い、30fpsの間隔でデータを記録した。

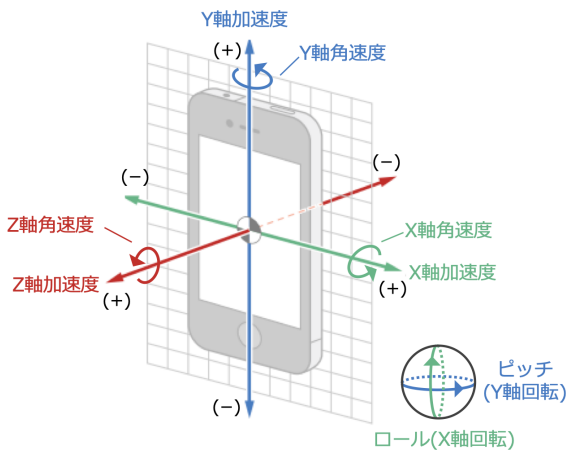


図2 スマートフォンの各センサの様子

データ収集タスクに使用したスマートフォンはサムスンのGalaxy S8+である。データ収集タスクのためのシステムの実装はJavaを用いて開発を行った。また、実験協力者は大学生の20代男性8名であり、それぞれの実験協力者がタスクについて1試行（漫画1冊分のタスク）とすることとした。

これらのデータ収集タスクにおいて、遅延が起きている遅延発生区間と、漫画を読んでいる読書区間について図3に示す。図のように、遅延があるページについては少し遅れてから読書が開始される。この遅延発生区間は実験協力者が待たされている状態にあり、実験協力者がストレスを感じやすい場面であると考えられる。これにより、遅延が発生した場面において収集したデータはストレスを感じている状態のデータと考えられ、遅延がない場面において収集したデータと異なり、特徴的な動きのデータとなっている可能性が考えられる。次章では、これらの遅延発生区間が含まれる場面のデータと遅延がない場面のデータを自動で判別できるかを実験によって明らかにし、ストレスを感じる場面を検出することが可能であるかについて検証する。

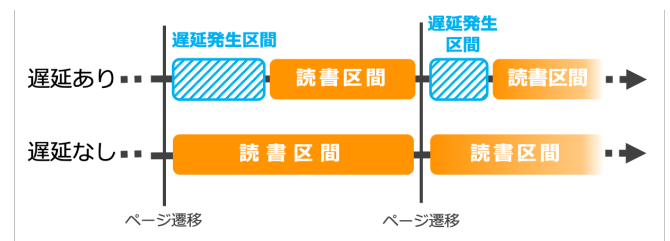


図3 遅延発生区間と読書区間

4. ストレス判定実験

前章で構築したデータセットを利用し、ユーザがストレスを感じている場面を自動で検出可能かについて実験を行う。本章では、機械学習によってデータからユーザのストレスを判定することを試みる。

4.1. 概要

ユーザがストレスを感じる場面を自動で検出するため、スマートフォンのセンサ情報から機械学習によってユーザのストレスを判定する。今回は、ユーザがストレスを感じている状態をストレス状態、それ以外の状態を非ストレス状態とし、それぞれの状態を判定するための2値分類の問題に取り組む。

まず、3章で収集したデータセットをもとに、データから特徴量の抽出およびデータごとにストレス状態・非ストレス状態のラベルの付与を行うことで、学習データの生成を行う。次に、それらの特徴を学習したモデルを作成し、データを実際に判定することでモデルの精度を検証する。

4.2. 学習データの生成

スマートフォンで取得したセンサ情報から、学習に利用する特徴量を生成する。データセットでは、加速度・角速度および姿勢についての情報を収集した。ここで、データセット構築でも述べたように、加速度と角速度についてはそれぞれX軸、Y軸、Z軸の3軸、

姿勢についてはロール、ピッチの2軸の情報がそれぞれ存在する。データはそれぞれの軸について値をもつため、合計8種類のデータが存在する。

本実験ではこれらのデータについて、遅延発生区間を含めた漫画1ページの読書におけるセンサ情報の最大値、最小値および平均値の3パターンの値を特徴量として生成した。8種類のデータそれぞれに3次元の値が生成されるため、合計24次元の特徴量を生成する。生成した特徴量をまとめたものを表1に示す。

表1 生成した特徴量と次元数

加速度	角速度	姿勢
X軸	X軸	ロール
Y軸	Y軸	ピッチ
Z軸	Z軸	

各特徴それぞれに最大・最小・平均の3次元が存在
合計 24次元

また、ストレス状態・非ストレス状態のラベルについては、遅延発生区間が含まれる読書場面のデータをストレス状態とする。実験協力者1人あたり最大50回の遅延があったため、それぞれの遅延発生のタイミングに該当するデータをストレス状態、それ以外の読書中のデータを非ストレス状態としてラベルの付与を行った。遅延発生のタイミングについては、次のページへ遷移を開始してから、ページが表示されるまでの区間としている。また、非ストレス状態のデータ数の方が多かったため、アンダーサンプリングによってそれぞれのラベルのデータ数を揃えた。

これらの結果、データセット構築の実験協力者8人分について合計698件（ストレス状態・非ストレス状態それぞれ349件）のデータが生成された。これらを学習データとして扱い、機械学習によるストレス判定実験に利用する。

4.3. 判定実験

学習データをもとに、ストレス状態・非ストレス状態を判定する2値分類の判定実験を行う。今回は分類問題の学習アルゴリズムとしてランダムフォレストを使用した。機械学習にはPythonのライブラリであるscikit-learn[11]を利用する。

判定実験では、データセット構築の実験協力者8人について、それぞれの個人のデータのみを利用してストレス状態・非ストレス状態を判定する実験と、全員のデータを利用して判定する実験を行う。全員のデータを利用した判定実験で作成される学習モデルは、どの実験協力者のデータであってもストレス状態・非ストレス状態を判定できるようになるため、ユーザ個人

に限定することなく汎用的に利用できる可能性が考えられる。

また、データ数が不足している可能性を考慮し、データセットの精度の検証には交差検定を利用する。今回は、データ数の4分の1をテストデータとして利用し、すべてのデータを重複なくテストデータとして扱う4分割交差検定(4-foldクロスバリデーション)を利用する。

これらの実験協力者各8人+全員のデータによる9種類の学習モデルについて、それぞれ交差検定の平均値をストレス判定の精度として求める。

4.4. 結果

実験協力者8人(A~H)および全員の判定実験における学習データの件数と精度を表2に示す。実験協力者ごとにデータ数が異なることについては、遅延発生の時間がランダムに設定されていたためである。

表2より、個人のデータのみを利用する場合は、平均65.1%の精度でストレス状態・非ストレス状態の判定が可能であり、精度の最大値は実験協力者Aの72.8%であった。また、全員のデータを利用する場合の判定精度は69.6%であり、個人のデータのみを利用する場合よりも比較的低い精度となることがわかった。

表2 判定実験の精度の比較

実験協力者	データ数	判定精度(%)
A	82	72.8
B	96	55.2
C	94	57.7
D	90	65.8
E	84	70.5
F	92	60.8
G	76	72.4
H	84	65.6
平均(A~H)	87	65.1
全員	698	69.1

また、全員のデータを利用した判定実験の結果について、漫画1ページを読むのにかかった時間(ページ滞在時間)および遅延発生区間の時間(遅延時間)の違いによる比較を行う。

表3は、ページ滞在時間について5秒刻みのグループに分け、各グループにおけるストレス状態のデータ数とその再現率を示したものである。表3の再現率は、判定実験の結果より、ストレス状態のデータを正しく

判別した割合となっている。

表 3 より、ストレス状態の全データ 349 件の再現率が 69.6%であるのに対して、10 秒～15 秒のグループにおける再現率が 80.9%、15 秒以上のグループにおける再現率が 92.1%と値が大きく上回る結果となった。これらの結果は、漫画 1 ページに対して長い時間をかけて読んでいる場面ほど、ストレス状態のデータを正しく判定することが可能であることを示している。

表 3 ページ滞在時間ごとの再現率

滞在時間	データ数	再現率
0 秒～5 秒	75	49.3%
5 秒～10 秒	168	69.0%
10 秒～15 秒	68	80.9%
15 秒以上	38	92.1%
全体	349	69.6%

また、表 4 は遅延時間について 2 秒刻みのグループに分け、各グループにおけるストレス状態のデータ数とその再現率を示したものである。

表 4 遅延時間ごとの再現率

遅延時間	データ数	再現率
0 秒～2 秒	138	51.4%
2 秒～4 秒	106	78.3%
4 秒～6 秒	64	84.4%
6 秒～8 秒	24	87.5%
8 秒～10 秒	17	82.4%
全体	349	69.6%

表 4 より、ストレス状態の全データ 349 件の再現率が 69.6%であるのに対して、遅延時間が 0 秒～2 秒のグループを除いたすべてのグループの再現率が上回る結果となった。特に、4 秒以上のグループは再現率が

80%以上となっている。これらの結果は、遅延時間が長いほど、ストレス状態のデータを正しく判定することが可能であることを示している。

また、ランダムフォレストは複数の決定木を用いて分類を行うアンサンブル学習の手法であるため、複数の決定木同士を比較することによって各特徴量の重要度を評価することができる。これにより、実験に利用した 24 次元の特徴量について重要度を評価する。また、特徴量は加速度 3 軸、角速度 3 軸、姿勢 2 軸の 8 種類のデータそれぞれについて、単位時間の最大値、最小値、平均値を計算している。そのため、8 種類のデータごとに最大値、最小値、平均値の重要度を合計し、データごとの重要度を評価する。なお、重要度の合計は 1.0 になるように正規化されている。実験協力者 8 人および全員のデータによる各学習における重要度の平均を求め、8 種類のデータの重要度を比較した結果を図 4 に示す。図 4 では、加速度、角速度、姿勢の情報ごとに色分けを行っている。

図 4 より、比較的重要度が高いのは X 軸の加速度および姿勢情報のロール(X 軸の回転方向)であったが、各特徴の重要度にそれほど大きな差は見られなかった。

5. 考察

判定実験の結果、個人のデータのみによる学習の精度は平均 65.1%であり、全員のデータによる学習の精度は 69.1%であることがわかった。これらの結果より、全員のデータを用いた汎用的な学習モデルであっても、個人のデータに限定した学習モデルと同様の精度となることがわかった。しかし、どちらの学習モデルについても判定精度が低く、ストレス状態・非ストレス状態を十分には判定できていないと考えられる。これについては、データセット構築の実験協力者も 8 人と少ないため、さらに人数を増やすことによってデータ数を増やす必要が考えられる。

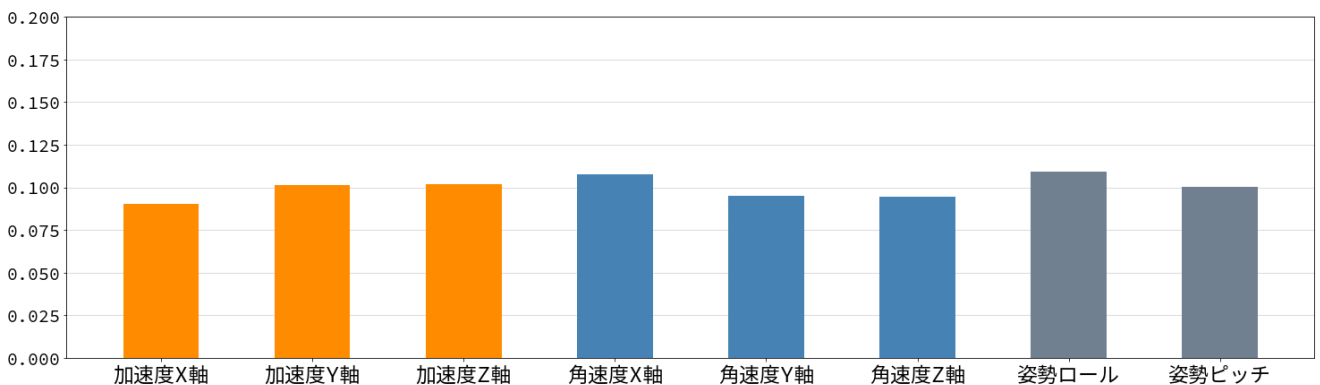


図 4 各特徴の重要度の比較

ここで、ページ滞在時間の違いによる再現率の比較を行ったところ、ページ滞在時間が長いほどストレス状態のデータの再現率が高いことがわかった。この結果より、漫画の1ページを長く鑑賞している場面ほどストレスを検出しやすい可能性を示している。これについては、滞在時間が長いことより、実験協力者が漫画を集中して読んでいる場面である可能性が考えられる。そのような場面において遅延によって焦らされることで、平時より大きなストレスを感じている可能性があり、判別が容易になったと考えられる。また、遅延時間が長いほど再現率が高いことより、長く待たされるほどストレス状態を判定できるという結果となった。しかし、これらのページ滞在時間および遅延時間について、時間が長いグループのデータが短いグループのデータより少ないため、再現率の差に影響を与えている可能性が考えられる。よって、このようなページ滞在時間・遅延時間が長い場合のデータを更に収集し、時間が短い場合のデータと同程度の数を集めることによって比較する必要が考えられる。

また、特徴量の重要度の評価では、突出して重要と判断された特徴は存在しなかった。これについては、学習の精度が低い原因の1つとして、有益な特徴を抽出しきれていない可能性が考えられる。よって、今後は時系列を考慮したデータの抽出など、様々な特徴抽出の方法を検討する必要がある。

また、判定実験の結果は2値分類における学習の精度としては低いと考えられるため、学習の方法やデータセット構築の改良が必要となる。今回の実験で精度が低かった理由としては、ストレス状態・非ストレス状態のラベルの付与が正確でなかった可能性が考えられる。今回は遅延が発生したタイミングのデータすべてをストレス状態のデータとして扱った。しかし、遅延が発生している区間に限定してユーザがストレスを感じているとは限らず、遅延後に継続してストレスを感じていることも考えられる。これらのデータのストレス状態・非ストレス状態のラベルの付与を正しく行うためには、ストレス状態であると判断する指標を並行して取得する必要がある。そのため、データセット構築時に、従来の研究で用いられている心拍数・発汗などの生理指標も同時に記録し、これらの情報からストレス状態を判断する必要がある。また、今回のデータセット構築におけるデータ収集タスクにおいて、タスクに対してストレスを感じない実験協力者が存在していた可能性も考えられる。特に、今回は漫画を読むタスクとしたが、普段から漫画を読まないなど、タスクに対して馴染みのない実験協力者は不満を抱かないことが考えられる。個人のデータのみによる推定実験の結果についても、判定精度が最も高い72.8%の実

験協力者に対し、精度が55.2%と低い実験協力者も存在していた。このように、個人によって不満に感じる度合いに差があり、精度に影響を及ぼしていると考えられる。そのため、今後はデータセット構築の実験協力者を増やすとともに、よりデータ収集タスクと同じ状況に馴染みのある実験協力者を集めることが課題となる。

6. プロGRESSバーによるストレス判定実験

判定実験の結果より、判定精度が低かった原因の1つとして、ストレス状態・非ストレス状態のラベルの付与が正確でなかったことが考えられる。そこで、ストレス状態・非ストレス状態をより正確に定義するための新しいデータセット構築を行い、判定実験を再度行う。新しいデータセット構築では、画面の読み込み時にPROGRESSバーを提示することで、実験協力者のストレスを制御する方法を試みる。

6.1. データ収集

3章のデータセット構築と同様に、漫画の読書中にページ読み込みによる遅延が発生するタスクを用意した。ここで新しいデータセット構築では、ページ読み込みにおいてストレスを感じにくい状況をつくるため、ページ読み込みの進捗を可視化したPROGRESSバーを提示することとする。PROGRESSバーは、情報の読み込みについてどの程度読み込みが完了したかを可視化したものであり、これらの提示はユーザが感じる待機時間を短く感じさせるという特徴がある[12]。これにより、遅延が発生している区間について提示があることがストレスの軽減に繋がると考えられる。よって、タスクにおいて遅延発生区間にPROGRESSバーを提示する場合と提示しない場合を作り、提示なしをストレス状態、提示ありを非ストレス状態としてラベルを付与することとする。



図5 PROGRESSバーの提示による違い

新しいデータ収集タスクでは、図5のようにPROGRESSバーの提示ありと提示なしの場面を用意する。ページの読み込みによる遅延は1冊の漫画に対して10

回発生し、プログレスバーの提示あり、提示なしがそれぞれ5回ずつ発生することとする。また、遅延の発生時間は5秒~14秒とする。ここで、前回のデータ収集タスクでは1冊の漫画に対し遅延が50回であったのに対し、新しいデータ収集タスクでは遅延を10回と少なくした。この理由については、200ページ前後の漫画に対して50回の遅延が起きるのは多すぎると判断したためである。

これらのタスクを、実験協力者である大学生の20代男性6名に対して行い、それぞれの実験協力者がタスクについて2試行(漫画2冊分のタスク)を行うこととした。

6.2. 判定実験

新しく収集したデータセットを利用して、再び判定実験を行う。4章の判定実験と同じく、ユーザがストレスを感じている状態をストレス状態、それ以外の状態を非ストレス状態とし、それぞれの状態を判定するための2値分類の問題に取り組む。

今回の判定実験では前節で述べたように、ストレス状態・非ストレス状態のラベルを遅延発生時のプログレスバーの有無によって付与する。よって、プログレスバーの提示なしのデータをストレス状態のデータ、プログレスバーの提示ありのデータを非ストレス状態のデータとした。また、特徴量については4章の判定実験と同じく、ストレス状態・非ストレス状態におけるセンサ情報について、遅延発生区間の平均値・最大値・最小値を特徴量とする。これらを学習データとし、データセット構築の実験協力者6人分について合計120件(ストレス状態・非ストレス状態それぞれ60件)のデータが生成された。

これらの学習データをもとに、ストレス状態・非ストレス状態を判定する2値分類の判定実験を行う。分類のアルゴリズムはランダムフォレストを使用し、4分割交差検定(4-foldクロスバリデーション)によって、交差検定の平均値をストレス判定の精度として求める。

6.3. 結果・考察

判定実験の結果、120件のデータに対する判定精度は55.0%と低い結果になった。また、120件のデータにおいて、遅延中にページの移動などの操作が行われたデータが21件存在した。これは、遅延が発生している間に前のページを見返すなど、読み込みが終わるまで他のページを読むために操作したものと考えられる。また、これらの操作はプログレスバーの提示ありで11件、提示なしで10件起こっていたため、判定におけるノイズになっていた可能性も考えられる。しかし、21件のデータを除いた99件によって学習および判定を行ったところ、精度は52.5%と低い結果となった。よ

って、今回の実験ではストレス状態・非ストレス状態を判定することが難しいという結果となった。

前回の判定実験の精度よりも低い原因の1つとしては、新しいデータセット構築において遅延発生回数を減らしたことにより、学習に利用するデータが少なくなってしまうことが考えられる。また、データ収集タスクにおいて、プログレスバーの提示によって待機時間を短くすることにより、実験協力者が感じるストレスをコントロールすることを試みたが、これらの提示がストレスに大きな影響を及ぼすことができなかったことが考えられる。

判定精度を上げるための今後の改善としては、心拍などの生体情報やデータ収集タスクの実行中の行動から定量的にストレスを計測し、より正確にストレスを感じた箇所を定義する必要があると考えられる。今回はプログレスバーの提示がない遅延発生区間がストレスを感じる場面と定義したが、プログレスバーの提示が実際にストレスの軽減に繋がったかについて十分な検証ができていないという問題がある。よって、ストレスを観測するために一般的に利用される心拍などの利用により、より明確にストレス状態を定義する必要があると考えられる。また、遅延発生区間の前後にもストレスを感じている可能性など、どのタイミングでストレスを感じているか曖昧であるという問題も存在する。これらについても、定量的なストレスの計測によってモニタリングすることで、どの区間において実験協力者がストレスを感じていたかを明確化することができると考えられる。よって今後は、定量的にストレスを計測する方法を組み入れ、ストレスをできるだけ正確に評価した上で判定実験を行う必要がある。

7. まとめ

本研究では、サービス開発者がサービスの問題点を自動的に検出することを可能にするために、スマートフォンのセンサ情報からユーザがストレスを感じている場面を検出する方法を検討した。そのため、実際のスマートフォンの利用状況を想定したタスクを用意し、ユーザのストレス時のセンサ情報をデータセットとして収集した。また、それらのデータをもとに機械学習によってストレス判定実験を行ったところ、個人のデータのみで学習で平均65.1%、全員のデータによる学習で69.1%の精度で判定できることがわかった。また、プログレスバーの提示の有無によってストレス状態を制御する新しいデータセットを収集し、判定実験を行ったが、精度は55.0%と低い結果となった。

今後はより厳密にストレス時のデータを収集するため、他のストレスを検出するための指標を利用し、より多くのデータ収集を行う。また、これらのストレ

ス判定を利用することで、実際のサービスの問題点をどの程度検出できるかについて調査を行っていく予定である。

謝辞 本研究の一部は、JST ACCEL（グラント番号 JPMJAC1602）の支援を受けたものである。

参 考 文 献

- [1] 中川千鶴. 特集③人間工学のための計測手法 第4部：生体電気現象その他の計測と解析（5）－自律神経系指標の計測と解析－. 人間工学, Vol. 52, No. 1, pp. 6-12, 2016.
- [2] J. Nielsen. Usability inspection methods. Conference Companion on Human Factors in Computing Systems, pp. 413-414, 1994.
- [3] 坂井誠, 中道上, 島和之, 中村匡秀, 松本健一. WebTracer: 視線を利用した Web ユーザビリティ評価環境. 情報処理学会論文誌, Vol. 44, No. 11, pp. 2575-2586, 2003.
- [4] F. Paternò, A. G. Schiavone, A. Conti. Customizable Automatic Detection of Bad Usability Smells in Mobile Accessed Web Applications. Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, pp. 1-11, 2017.
- [5] R. Eardley, A. Roudaut, S. Gill, S. J. Thompson. Understanding Grip Shifts: How Form Factors Impact Hand Movements on Mobile Phones. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 2017.
- [6] Kazuki Tajima, Satoshi Nakamura. WePatch: A System Enabling Users to Improve Bad User Interfaces on the Web, OzCHI 2017, pp. 448-451, November 2017, Brisbane, QLD, Australia.
- [7] 小川剛史, 佐藤博則, 狩川大輔, 高橋信. ライフログデバイスをを用いた複雑タスク環境下におけるワークロード推定に関する研究. ヒューマンインタフェース学会論文誌, Vol. 19, No. 4, 2017.
- [8] JINS MEME|TURN IT ON - 見るから、知るへ、(最終閲覧日 2019 年 1 月 10 日) <https://jins-meme.com/ja/>
- [9] J. Hernandez, P. Paredes, A. Roseway, M. Czerwinski. Under pressure: sensing stress of computer users. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2014.
- [10] A. Fujimoto, T. Ogawa, K. Yamamoto, Y. Matsui, T. Yamasaki, K. Aizawa. Mangal09 dataset and creation of metadata. Proceedings of the 1st International Workshop on coMics ANalysis Processing and Understanding, pp. 2:1-2:5, 2016.
- [11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay. Scikit-learn: Machine Learning in Python, The Journal of Machine Learning Research, pp. 2825-2830, 2011.
- [12] 松井啓司, 中村聡史, 鈴木智絵, 山中祥太. 周辺視野への視覚刺激提示によるプログレスバーの主観的な待機時間短縮手法. 情報処理学会 研究報告ヒューマンコンピュータインタラクション (HCI), 2019-HCI-181 (25), 1-6 (2019-01-14).