

# 小型ウェアラブルカメラを用いた プライバシーウェアな行動認識の実現に向けた基礎調査

平岡 滉司<sup>†1,a)</sup> 中村 優吾<sup>†2,b)</sup> 荒川 豊<sup>†2,c)</sup>

**概要:** 加速度センサを用いた行動認識は、低コスト・低消費電力であるという利点から、これまで多くの研究が行われてきた。しかし、より多様な行動を対象とし、堅牢かつ汎化性能の高い行動認識を実現するためには多くの課題が残っている。この課題を克服するために、より詳細な視覚的特徴を観測可能なウェアラブルカメラを活用するアプローチが有望視されている。一方で、カメラ映像には人の顔やプライバシーに関わる情報が映り込む可能性があるため、ユーザーの受容性を低下させる懸念がある。本研究では、加速度データに加えて、プライバシー処理を施した手元の画像データを用いることで、認識精度の向上とユーザーのプライバシーに配慮した行動認識の両立を目的とする。行動認識の一例として、デスクワーク6種(タイピング、マウス操作、スマートフォン操作、飲む、書く、その他)を対象とし、手首に装着したウェアラブルカメラから取得した、加速度、画像、画像(プライバシー処理済み)、加速度+画像、加速度+画像(プライバシー処理済み)について認識精度を比較し、提案手法の有効性を評価した。

## 1. はじめに

加速度センサを用いた行動認識は、低コスト・低消費電力であるという特性から、ウェアラブルデバイスにおける基盤技術として広く採用されている。この技術は、ヘルスケア、スポーツ、介護支援といった多種多様な分野に応用され、その有効性が確認されている。しかし、加速度センサのデータのみでは、特定の行動の背後にあるコンテキストを完全には把握できず、類似した動作の分類や異なるユーザー間での汎化性能を有する推論モデルの構築が困難であるという問題が存在する [1]。この問題に対処するために、異なるセンサを組み合わせるアプローチや、複数の場所に加速度センサを装着する方法が提案されているが、これだけでは不十分である場合が多い [2]。

そこで、加速度センサよりも豊富な情報を提供するウェアラブルカメラを使用したアプローチが有望視されている [1]。カメラから得られる視覚情報は、色や背景情報など、加速度センサが十分に取得できなかった情報を補完し、よりロバストな行動認識システムの構築が期待される [3]。しかし、視覚情報の利用には、プライバシー保護の観点からさまざまな課題が存在する。例えば、カメラ映像には顔

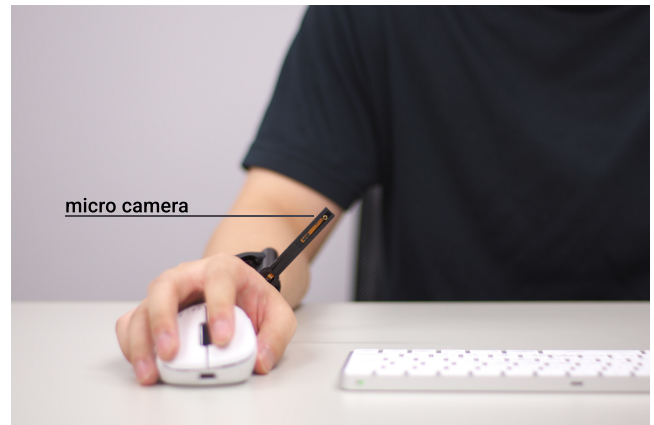


図 1: 小型ウェアラブルカメラ装着時の様子

や個人を識別可能な情報が含まれる可能性があり、ユーザーやその周囲の人々に不安や不快感を与える原因となる。これらの課題が解消されない限り、ウェアラブルカメラによるライフログや日常生活動作 (ADL) のセンシングにおいて、適用範囲が限定的になり、実用化は難しい。このため、ウェアラブルカメラの実用化に向けて、プライバシー保護のためのぼかし処理、データの匿名化といった手法が検討されており、ユーザーのプライバシーを保護しつつ高精度な行動認識を実現する方法が模索されている [4]。

本研究では、ユーザーのプライバシー保護と行動認識精度の向上を両立させた行動認識の実現を目指し、加速度データに加えて、プライバシー処理を施した画像データを活用した行動認識手法に関する基礎評価を行った。画像の

<sup>†1</sup> 九州大学大学院システム情報科学府

<sup>†2</sup> 九州大学大学院システム情報科学研究院

a) hiraoka.koushi.031@s.kyushu-u.ac.jp

b) y-nakamura@ait.kyushu-u.ac.jp

c) arakawa@ait.kyushu-u.ac.jp

プライバシーを保護する具体的な手法として、ガウスぼかし処理 (GaussianBlur)、アニメ化 (CartoonGAN)、エッジ検出 (Canny Edge) という3つを採用した。ガウスぼかし処理では、顔やディスプレイ情報など、個人識別可能な情報を含む映像に対してぼかしのレベルを調節し、個人情報の保護を試みる。アニメ化処理では、画像をアニメスタイルに変換し、顔などの具体的な識別情報を隠蔽しつつ大まかな特徴を保持することで、プライバシー保護と情報の有用性をバランスさせる。エッジ検出処理は、画像をモノクロ化し、物体の輪郭のみを抽出することで、より高いプライバシー保護が期待できる手法である。

評価シナリオには、手首に装着されたウェアラブルカメラを使用し、デスクワークにおける様々な行動 (タイピング、マウス操作、スマートフォン操作、飲む、書く、その他) を認識する場面を採用し、それぞれの手法のプライバシー保護効果および行動認識への影響について基礎評価を行った。プライバシー保護効果は、それぞれのプライバシー保護処理を施した映像データを用いてプライバシーに関するアンケート調査を行い、その差を  $t$  検定によって評価した。また、行動認識への影響については、加速度データ、画像データ、プライバシー処理済み画像データ、およびそれらの組み合わせを入力データとして機械学習モデルを構築し、デスクワーク行動の認識精度を評価した。

評価実験では、9人の被験者がウェアラブルカメラを装着し、10分間5種のデスクワークを実施した。また、オンラインアンケートにより、30人の被験者が参加し、プライバシー処理された各10秒程度の動画を視聴後、プライバシーに関する6つの質問に回答した。

アンケート調査では、プライバシー保護処理によって、行動認識におけるカメラ利用時の抵抗感を低減させることが明らかになった。画像利用に対する不安度はエッジ検出が最も低く、対応あり  $t$  検定を用いてエッジ検出と他のプライバシー保護処理方法を比較した結果、全ての検定において  $p < 0.05$  であり、デバイス受容性においてエッジ検出が効果的であることが示された。

行動認識では、加速度データと画像データの組み合わせによって、それぞれ単体の利用時に比べ認識精度が向上した。特に、入力データとして、加速度データとエッジ検出を組み合わせるとき、最も高い認識精度を示し、単体の加速度データと比較して20.3%、単体の画像データと比較して4.1%の精度向上を達成した。

したがって、Canny Edge Detection を用いたエッジ検出と加速度データを組み合わせた行動認識は、本研究の比較対象の中で、定性的なプライバシー評価、定量的な認識精度評価において最も優れた手法であり、プライバシーウェアな行動認識の実現に向けて有望な選択肢であることを示した。

## 2. 関連研究

本章では、行動認識で主流である加速度センサやウェアラブルカメラとともに、複数のセンサを組み合わせた手法を述べる。また、画像のセンシング利用における個人プライバシーへの対応について言及する。

### 2.1 ウェアラブルデバイスを用いた行動認識

#### 2.1.1 加速度センサベース

加速度センサは、そのコンパクトさ、低消費電力、低コストといった特性から、センサベースの行動認識において広く研究が行われてきた [5]。その応用範囲は広がっており、スマートフォンやスマートウォッチなど、さまざまなウェアラブルデバイスに組み込まれていることから、日常的な行動を認識するための基盤技術である [6]。しかし、同一の動きであっても、コンテキストの理解には情報量が十分でなく、異なる行動として解釈されることがある [7]。このため、加速度センサ単独の行動認識には限界があることが指摘されている [1, 8]。

#### 2.1.2 カメラベース

カメラによる視覚情報は、色や背景などシーン全体のコンテキストの理解に役立つ。エゴセントリック (自己視点) のカメラ映像では、Aria [9] や Ego4D [10] が世界各地で収集された膨大な量の映像データセットを提供している。このデータセットはアノテーション付きであり、日常生活動作 (ADL) の解析に利用されている [11]。

より詳細な行動の認識のために、手首に装着されたリストバンド型のウェアラブルカメラ [12, 13] が提案されている。ADLの認識のために、様々な日常行動において、他のエゴセントリック視点の装着位置と比較し、手首視点のカメラの有効性を示した [13]。また、手のジェスチャー認識において、手首視点のカメラの利用は進展しており、家庭用ロボットや車椅子操作のために、1350のジェスチャーサンプルを分類し、平均認識精度97.6%を達成している [12]。

#### 2.1.3 マルチセンサベース

単一加速度センサのコンテキスト理解の限界を補うために、カメラを組み合わせることで認識精度の向上と分類可能な行動の多様化をねらう。音声 [14] や無線 [15]、その他センサも行動認識システムで活用されているが、ウェアラブルデバイスとして装着に不向きであったり、ノイズの影響を受けやすいという制約がある [16]。そのため、加速度センサやカメラに比べて適用範囲やシナリオが限定されている場合がある。スマートフォンやスマートグラスには多くの場合、加速度センサが内蔵されている。その特性を活かして、カメラと組み合わせた行動認識が盛んに行われてきた [17, 18]。また、コンパクトなセンサであることから、ウェアラブルデバイスとしての実装が容易であり、ヘッド

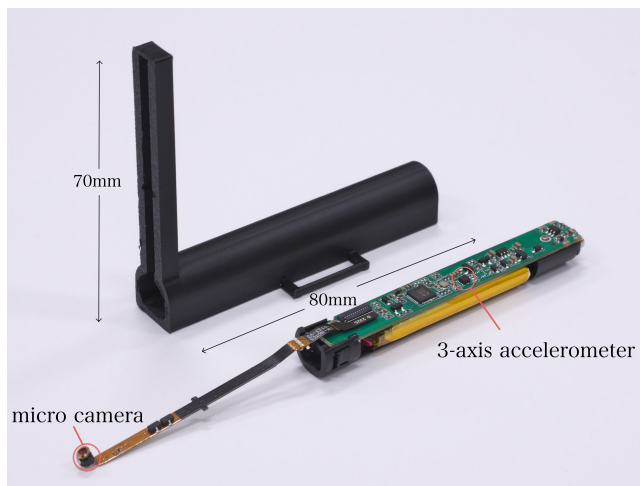


図 2: リストバンド型ウェアラブルカメラ

セット型 [10,19] やボディウェア型 [20,21], リストバンド型 [3,12] など, 様々な形式のウェアラブルデバイスが提案されている。

## 2.2 センシングにおけるカメラ利用のプライバシー問題

カメラから取得する映像データには必要以上に詳細な情報が含まれているため, 個人プライバシーの侵害が想定される。実環境において, カメラを用いたセンシングはユーザ自身や周辺の人々に不安や不快感を与え, 潜在的に行動が変化したり, デバイスの着用を放棄したりする危険性を指摘している [4,22]。

そこで, プライバシーに配慮したカメラのセンシング利用のために, 顔認識システムを組み合わせたマスキング手法 [23] やぼかし技術 [24] が提案されている。ウェアラブルカメラから取得される画像にもプライバシー処理が検討されており, 手元の動作に関連する 10 種類の行動を収集し, ぼかしやエッジ検出, アニメ化やアバターによるマスキングなど, 異なる画像処理によるプライバシーに関するアンケート調査を実施した。その結果, アニメ化による画像処理が個人の特徴を効果的に隠す手法として高く評価された [22]。

本研究では, リストバンド型ウェアラブルカメラを用いて, 加速度センサでは不十分な情報をカメラで補うことで, 行動認識の精度向上をねらう。また, 日常利用において, 手元の画像利用はデバイスの画面や人の顔が映り込む危険性から個人プライバシーの問題が発生する。そのため, 取得映像の画像処理によってユーザのプライバシーに配慮し, 定性評価を実施することで, プライバシーアウェアな行動認識システムを構築する。

## 3. 提案手法

### 3.1 使用したデバイス

本研究では, 所属研究室で開発したリストバンド型の小

型ウェアラブルカメラ (図 2) を使用した。このデバイスには, 3 軸加速度センサとマイクロカメラが内蔵されており, 加速度データはサンプリングレート平均 25Hz, 画像データはフレームレート最大 15fps で記録される。加速度データに加え, 単一デバイスから取得した手元の画像データを活用し, 動きにオブジェクト情報を補完することで, より詳細な行動認識の実現を目指す。

### 3.2 データ収集

リストバンド型ウェアラブルカメラを用いた行動認識の一例として, オフィス環境におけるデスクワークを対象とした。日常のデスクワークにおいて, 手の動きを伴う 5 つの行動 (図 3: タイピング, マウス操作, スマートフォン操作, 書く, 飲む) に加え, 対象行動以外の動き (その他) を含めた 6 クラスを分類する。データ収集は, 実験参加者 9 名の学生に対して行い, 右手首にウェアラブルカメラを装着 (図 2) し, 5 つの対象行動を各 2 分間, 計 10 分間測定した。対象行動以外の動き (その他) を収録するために, 準備時間や前後の動きのない時間も含めて測定した。

### 3.3 画像のプライバシー処理

カメラを用いたセンシングでは, ユーザが意図せず顔やプライバシー情報を収集できることが問題点として挙げられる。そこで, 以下の画像処理手法を用いてプライバシーに配慮した行動認識を目指す (図 4)。

#### 3.3.1 GaussianBlur (ぼかし処理)

GaussianBlur は, ガウス分布に基づいたぼかし処理を適用する。ksize(kernel size) によって畳み込みサイズを調整することができる。従来研究では, プライバシーを考慮した屋内監視に ksize=30 前後が用いられている [25]。本研究では, 手元の画像を収集しているため顔や PC 画面などのプライバシーに関する情報に近接している場面が多い。そのため, ksize=30 程度では, 個人の顔を特定することが可能であるため, 経験則より顔や PC 画面の内容を認識することができないサイズ (ksize=77) にまで調整し, 学習に用いた。

#### 3.3.2 CartoonGAN (アニメ化)

CartoonGAN [26] では, 2 つの CNN から構成された GAN フレームワークを用いて実世界の画像をアニメイラスト調に変換することができる。アニメ化によって画像の大まかな特徴を残しながら詳細な特徴を隠すことができる。

#### 3.3.3 Canny Edge (エッジ検出)

Canny Edge Detection (Canny 法) によるエッジ検出では, 画像をモノクロ化することで背景情報を完全にフィルタリングでき, 物体の外形を検出する際に高いレベルのプライバシー処理が可能である [27,28]。ウェアラブルカメラは, 加速度センサに比べて消費電力が高いものの, グレースケール化によって画像サイズを大幅に圧縮できるため,



図 3: デスクワークにおける 5 つの行動の様子: (a) タイピング, (b) マウス操作, (c) スマートフォン操作, (d) 書く, (e) 飲む

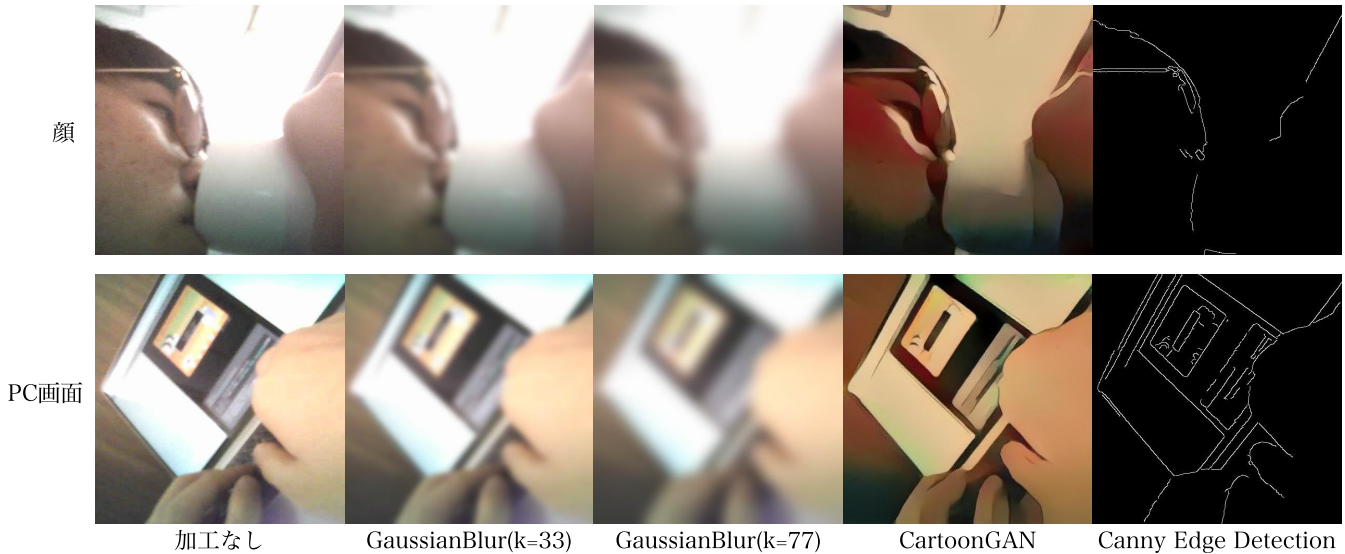


図 4: プライバシー保護のための加工処理が施された画像データ

比較的消費電力が低いセンシングも期待される。

### 3.4 システム構成

本章では、リストバンド型ウェアラブルカメラを用いた行動認識の提案システムについて述べる(図 5)。

#### 3.4.1 前処理

加速度はデータの欠損部分を含めてサンプリングレート平均 25Hz で記録される。前処理として、1 秒ごとに線形補間し、サンプリングレート 32Hz に揃えた。

また、画像サイズは全て  $224 \times 224$  にリサイズし、加速度+画像データでの学習時、特徴量間のスケールの違いを取り除くため、加速度データと画像データの特徴量に対して標準化を行った。

#### 3.4.2 特徴量抽出

加速度データと画像データはそれぞれウィンドウサイズ 2[s], スライディングウィンドウ 1[s] で特徴量抽出した。加速度データの特徴量として、データの平均・最大値・最小値・分散・最大周波数を用いた。画像データは、事前学習済みモデルである VGG16 の全結合層 3 層を除く 13 層の畳み込み層を用いて特徴量を抽出した。

ここで、加速度の特徴量ベクトルは 145 次元、画像の特徴量ベクトルは 4096 次元となる。加速度と画像の両方を入力に用いる場合、画像の特徴量ベクトルの次元数が加速度に比べて多いため、結合の際に互いの特徴を保つように画像から得られた特徴量を PCA (主成分分析) を用い、データの本質的情報を保持しながら次元削減した。

このとき、加工なしの画像の特徴量ベクトルをさまざまな次元数 (2, 3, 10, 50, 100, 145, 645, 1145, 1645, 2145, 2645, 3145, 3645, 4096) で試行し、加速度+画像データの認識精度を比較すると、145 次元の認識精度が最良であった(図 6)。したがって、提案システムでは画像の特徴量ベクトルを 145 次元で扱った。

#### 3.4.3 推定器と学習方法

本研究では、加速度データ、画像データ、加速度+画像データを入力として学習し、分類器として、以下の 3 つの教師あり機械学習モデルを選択した。

- LightGBM(LGBM)
- RandomForest(RF)
- SupportVectorMachine(SVM)

システムの性能評価のために、LOPO(Leave One Person

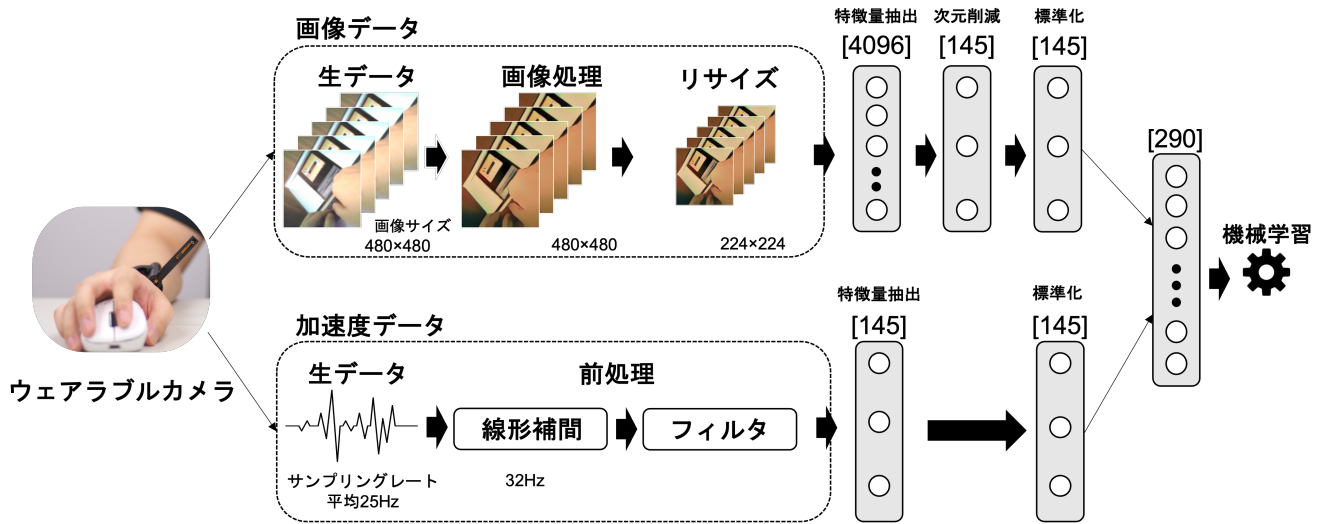


図 5: システム構成図

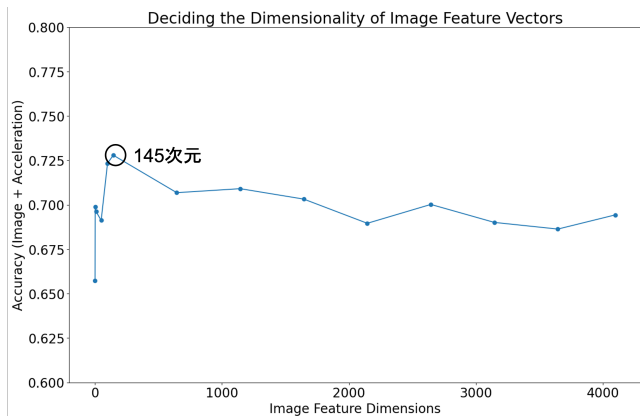


図 6: 各画像の次元数と加速度+画像データの認識精度 (分類器: RandomForest)

Out) 交差検証を用いて各入力データにおける最適な学習モデルを調査した。

#### 4. 評価実験

本研究では、手首に装着されたウェアラブルカメラを使用しデスクワーク行動認識シナリオにおいて、それぞれの手法のプライバシー保護効果および行動認識への影響について基礎評価を行った。4.1 節では、各プライバシー保護手法がユーザーのプライバシー意識に与える影響を評価した実験結果を、4.2 節では、各入力データを用いた行動認識精度を比較した実験結果を示す。

##### 4.1 リストバンド型ウェアラブルカメラ利用シーンにおけるプライバシーへの意識調査

ウェアラブルカメラの利用は、他のセンサに比べプライバシー侵害のリスクがある。行動認識におけるウェアラブルカメラの実用化に向けて、リストバンド型ウェアラブルカメラで記録された動画をもとに、ユーザーアンケートによ

るプライバシーに関する定性評価を実施した。

所属研究室メンバー 30 人を対象に、プライバシー処理した画像利用シーンによるリストバンド型ウェアラブルカメラの受容性を Fernandes らが作成した Qualtrics 調査 [22] に基づいて質問内容を改変し、アンケート調査を実施した。アンケート参加者は、デスクワーク監視時の 5 つの動画 (図 4) を視聴後、各動画について以下の質問に回答した。

アンケート参加者は各質問に 5 段階評価 (1: 全く不安に思わない, 5: かなり不安に思う) で回答し、その結果を集計した (表 1,2)。各動画における回答結果を「不安スコア」として定義し、プライバシーアウェアな画像処理について比較検討する。

ユーザ自身がリストバンド型ウェアラブルカメラを装着し、カメラによってセンシングされた場合の不安スコアを測定した結果 (Q2) によると、「加工なし」の条件では、不安スコアの平均値、中央値ともに最も高い数値である。実環境のセンシングにおいて、プライバシー処理を施さない画像利用はユーザの心理的負担が大きい。「ぼかし加工」は、平均値や中央値は「加工なし」と比較して不安スコアが改善されているが、前述 (3.3 節) 通り、接近して顔や個人のデバイスをセンシングしているため、プライバシー処理が不十分であり、スコアの分布は「加工なし」と類似している (図 7)。「ぼかし加工 (強)」について、個人の顔が特定できないレベルの強度なぼかし加工を加えた場合、「加工なし」の場合と比較して、画像利用への抵抗感は小さい。また、「アニメ化」は原型の特徴を保ちつつも、強度なぼかし加工と同様の成果を得られた。しかし、中央値は 5 段階中 4.0 である通り、詳細な行動は掴めないとしても背景情報や色情報は残されたままであり、実環境でのセンシングにおけるユーザへの心理的不快感は拭いきれない。これに対し、「エッジ検出」によるプライバシー処理では、不安スコアの平均値、中央値ともに最も低く、プライバシー保

護に対して有効であるといえる。これは、視覚的情報が大幅に簡略化されるため、他のプライバシー処理と比較してユーザに与える不安が軽減されると考えられる。

ユーザアンケート内容

- Q1** 添付されている動画にはリストバンド型ウェアラブルカメラで撮影した3つのシーンが含まれています。それぞれのシーンがユーザのどのような行動を撮影したものなのか簡潔に教えてください。もしわからない場合は「わかりません」と書いてください。
- Q2** あなたの手首にはウェアラブルカメラが装着されています。そのカメラで撮影された動画にあなた自身やあなたの持ち物が撮影されていた場合、プライバシー保護の観点でどの程度不安を感じるか評価してください。
- Q3-6** あなたの手首にはウェアラブルカメラが装着されています。そのカメラで撮影された動画に[家族(Q3)、友人(Q4)、同僚/知人(Q5)、見知らぬ人(Q6)]やその持ち物が映り込んでいた場合、プライバシー保護の観点でどの程度不安を感じるか評価してください。

表 1: ユーザアンケートの平均スコア

(1: 全く不安に思わない, 5: かなり不安に思う)

| 視聴動画         | Q2          | Q3          | Q4          | Q5          | Q6          |
|--------------|-------------|-------------|-------------|-------------|-------------|
| 加工なし         | 4.33        | 4.50        | 4.37        | 4.27        | 3.43        |
| ぼかし加工        | 3.90        | 4.10        | 3.93        | 3.83        | 2.90        |
| ぼかし加工(強)     | 3.50        | 3.60        | 3.40        | 3.33        | 2.43        |
| アニメ化         | 3.47        | 3.77        | 3.53        | 3.37        | 2.67        |
| <b>エッジ検出</b> | <b>2.57</b> | <b>2.83</b> | <b>2.60</b> | <b>2.50</b> | <b>1.97</b> |

表 2: Q2 のアンケートスコアの比較

| 視聴動画         | 平均値         | 中央値        | 標準偏差 |
|--------------|-------------|------------|------|
| 加工なし         | 4.33        | 5.0        | 1.03 |
| ぼかし加工        | 3.90        | 4.0        | 1.32 |
| ぼかし加工(強)     | 3.50        | 4.0        | 1.31 |
| アニメ化         | 3.47        | 4.0        | 1.38 |
| <b>エッジ検出</b> | <b>2.57</b> | <b>2.0</b> | 1.36 |

また、対応あり t 検定を用いて「エッジ検出」と他のプライバシー処理方法の不安スコアを検証した結果、全ての検定において  $p < 0.05$  であった。したがって、エッジ検出処理が他のプライバシー処理方法と比較して、リストバンド型ウェアラブルカメラの受容性を高められると推測される。

#### 4.2 各入力データにおける行動認識精度比較

リストバンド型ウェアラブルカメラから取得した加速度データ、画像データ、プライバシー処理済み画像データ、

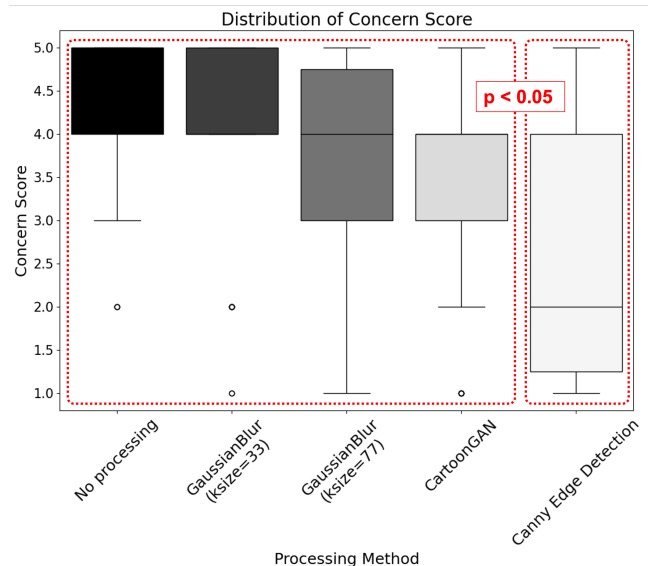


図 7: Q2 の回答結果による箱ひげ図

およびそれらの組み合わせを入力データとして、LGBM, RF, SVM によって機械学習モデルを構築した(表 3)。

学習の結果、全ての入力データにおいて、LGBM がその他 2 つの機械学習モデルと比較して認識精度を上回った。

#### 4.2.1 単一データ(加速度データおよび画像データ)による認識精度

加速度データ単体の認識精度は全体的に低く、いずれも 50% 台である。これに対し、画像データ単体では大幅に認識精度が向上している。本研究の実験シナリオにおいて、画像データの利用が有用であることを示している。また、プライバシー処理した画像において、「ぼかし加工(強)」や「アニメ化」に対して認識精度が大きく低下したのに対し、「ぼかし加工」や「エッジ検出」では「加工なし」と比較して 2~3% の低下に留まっている。ここで、顔や背景情報が保持されている「ぼかし加工」に対し、「エッジ検出」は、色や背景情報が損失しているのにも関わらず認識精度の劣化を抑制できていることから、対象オブジェクトのエッジ情報が行動認識において重要な情報であり、色や背景情報の損失の影響が小さいことを示唆している。

#### 4.2.2 画像データの特徴量における次元削減の影響

3.4.2 節より、画像の特徴量を 4096 次元から 145 次元まで削減したため、画像の単体利用において次元削減の影響を調査した。SVM において、次元削減による認識精度が著しく低下している。一方、LGBM や RF では次元削減によって精度が向上している。このことから次元削減によって冗長な特徴が削減され、行動認識に有用な情報を保持していることを示している。

#### 4.2.3 加速度データと画像データの組み合わせによる認識精度の変化

加速度データと画像データを組み合わせることで、全体的に認識精度が向上した。特に、「加工なし」では、LGBM

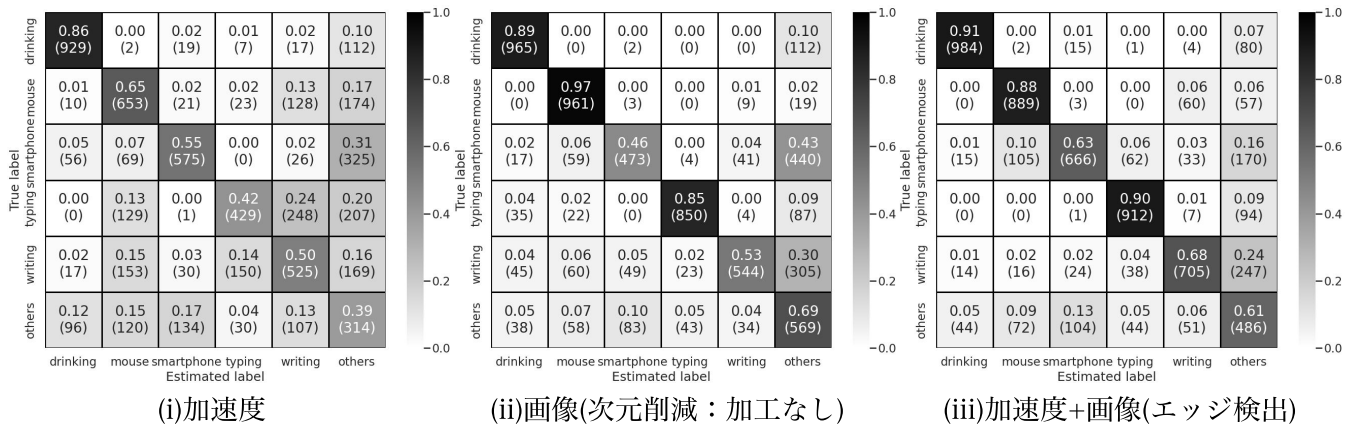


図 8: (i) 加速度データ, (ii) 画像データ (次元削減:加工なし), (iii) 加速度+画像データ (エッジ検出) の混同行列

表 3: 各入力データの認識精度比較

| 入力データ            | 機械学習モデル     |             |             |
|------------------|-------------|-------------|-------------|
|                  | LGBM        | RF          | SVM         |
| 加速度              | 57.0        | 53.5        | 52.7        |
| 画像 (加工なし)        | 72.0        | 69.4        | <b>68.9</b> |
| 画像 (ぼかし加工)       | 65.9        | 56.8        | 57.2        |
| 画像 (ぼかし加工強)      | 61.8        | 57.9        | 58.1        |
| 画像 (アニメ化)        | 65.6        | 62.9        | 62.7        |
| 画像 (エッジ検出)       | 69.0        | 62.9        | 65.8        |
| 画像 (次元削減:加工なし)   | 73.2        | 72.1        | 38.5        |
| 画像 (次元削減:ぼかし加工)  | 71.9        | 62.3        | 35.8        |
| 画像 (次元削減:ぼかし加工強) | 61.2        | 61.8        | 31.3        |
| 画像 (次元削減:アニメ化)   | 68.7        | 62.8        | 38.4        |
| 画像 (次元削減:エッジ検出)  | 70.2        | 65.7        | 36.7        |
| 加速度+画像 (加工なし)    | 76.7        | <b>72.7</b> | 57.4        |
| 加速度+画像 (ぼかし加工)   | 75.5        | 67.2        | 56.6        |
| 加速度+画像 (ぼかし加工強)  | 70.7        | 66.9        | 55.2        |
| 加速度+画像 (アニメ化)    | 71.4        | 64.0        | 56.2        |
| 加速度+画像 (エッジ検出)   | <b>77.3</b> | 71.1        | 56.0        |

で加速度データ単体と比較して 19.7%, 画像データ単体と比較して 3.5%の精度向上を達成した。また, プライバシー処理画像において, 「エッジ検出」が「加工なし」の認識精度を LGBM において 0.6%上回っており, 全体の中で最も認識精度の高い結果となった。「スマートフォン操作」, 「タイピング」, 「書く」など細かな動作が含まれる行動に場合であっても, 従来のセンシング手法に比べ大幅に認識精度が改善されている (図 8)。したがって, 情報量の限られたエッジ検出による画像データが加速度データと相補的に機能し, 精度向上に寄与している可能性を示唆している。

## 5. おわりに

本研究では, ユーザーのプライバシー保護と行動認識精度の向上を両立させた行動認識の実現を目指し, ウェアラブルカメラを用いたデスクワーク行動認識というシナリオを想定し, 加速度データとプライバシー処理を施した画像

データを活用したプライバシーウェアな行動認識手法に関する基礎評価を行った。4.1, 4.2 節の結果から, 画像利用では, 「エッジ検出」によるプライバシー処理が他の処理手法に比べ, 有意にユーザーの不安スコアを改善する結果となった。また, 加速度+画像データ (エッジ検出) において, 全ての入力データの中で最も高い認識精度 77.3%を示し, 単体の加速度データと比較して 20.3%, 単体の画像データと比較して 4.1%向上した。

このことから, ガウスぼかし処理 (GaussianBlur), アニメ化 (CartoonGAN) というプライバシー保護処理を適用した手法と比較して, Canny Edge Detection を用いたエッジ検出と加速度データを組み合わせた行動認識は, 定性的なプライバシー評価, 定量的な認識精度評価において最も優れた手法であり, プライバシーウェアな行動認識の実現に向けて有望な選択肢であることを示した。ただし, プライバシーは主観的価値に基づくため, 多様なユーザーにとって, この手法が十分なプライバシー保護を実現しているのか, また他のシナリオや条件下でも同様に有効であるのかについては未だ明らかではない。そのため, さらなる調査が必要である。

今後は, より細かな行動の認識や, 行動の質自体を評価する認識タスクに対する適応可能性を検討し, 家庭内や介護支援現場での行動モニタリングといったより高いレベルのプライバシー保護が求められるシーンにおいて, 提案手法がどの程度有効であるかを検証する。また, 計算資源が限られたウェアラブルデバイス上で, プライバシーウェアな行動認識タスクをオンデバイスで実行できるよう, 認識精度やプライバシー保護度合いに加えて, 計算負荷やメモリ容量の観点からも提案手法の改良および評価を進める予定である。

**謝辞** 本研究の一部は, JST さきがけ (JPMJPR21P7) の助成を受けたものである。

## 参考文献

- [1] Tong, C., Taylor, S. A. and Lane, N. D.: Are accelerometers for activity recognition a dead-end?, *Proceedings of the 21st international workshop on mobile computing systems and applications*, pp. 39–44 (2020).
- [2] Li, J., Umezawa, T. and Osawa, N.: Daily Activity Recognition Based on Acceleration of Both Wrists, *Human-Computer Interaction. Interaction Technologies: 20th International Conference, HCI International 2018, Las Vegas, NV, USA, July 15–20, 2018, Proceedings, Part III 20*, Springer, pp. 306–320 (2018).
- [3] Chen, F., Lv, H., Pang, Z., Zhang, J., Hou, Y., Gu, Y., Yang, H. and Yang, G.: WristCam: A wearable sensor for hand trajectory gesture recognition and intelligent human–robot interaction, *IEEE Sensors Journal*, Vol. 19, No. 19, pp. 8441–8451 (2018).
- [4] Alharbi, R., Tolba, M., Petit, L. C., Hester, J. and Alshurafa, N.: To mask or not to mask? balancing privacy with visual confirmation utility in activity-oriented wearable cameras, *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, Vol. 3, No. 3, pp. 1–29 (2019).
- [5] Bulling, A., Blanke, U. and Schiele, B.: A tutorial on human activity recognition using body-worn inertial sensors, *ACM Computing Surveys (CSUR)*, Vol. 46, No. 3, pp. 1–33 (2014).
- [6] Lara, O. D. and Labrador, M. A.: A survey on human activity recognition using wearable sensors, *IEEE communications surveys & tutorials*, Vol. 15, No. 3, pp. 1192–1209 (2012).
- [7] 井上創造: ウェアラブルセンサを用いたヒューマンセンシング, 知能と情報, Vol. 28, No. 6, pp. 170–186 (2016).
- [8] Lockhart, J. W. and Weiss, G. M.: Limitations with activity recognition methodology & data sets, *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, pp. 747–756 (2014).
- [9] Lv, Z., Charron, N., Moulon, P., Gamino, A., Peng, C., Sweeney, C., Miller, E., Tang, H., Meissner, J., Dong, J. et al.: Aria Everyday Activities Dataset, *arXiv preprint arXiv:2402.13349* (2024).
- [10] Grauman, K., Westbury, A., Byrne, E., Chavis, Z., Furnari, A., Girdhar, R., Hamburger, J., Jiang, H., Liu, M., Liu, X. et al.: Ego4d: Around the world in 3,000 hours of egocentric video, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18995–19012 (2022).
- [11] Liu, M., Ma, L., Somasundaram, K., Li, Y., Grauman, K., Rehg, J. M. and Li, C.: Egocentric activity recognition and localization on a 3d map, *European Conference on Computer Vision*, Springer, pp. 621–638 (2022).
- [12] Nguyen, H.-Q., Le, T.-H., Tran, T.-K., Tran, H.-N., Tran, T.-H., Le, T.-L., Vu, H., Pham, C., Nguyen, T. P. and Nguyen, H. T.: Hand Gesture Recognition From Wrist-Worn Camera for Human–Machine Interaction, *IEEE Access*, Vol. 11, pp. 53262–53274 (2023).
- [13] Kakaraparthi, V., Goldberg, M. and McDaniel, T.: Wrist view: understanding human activity through the hand, *International Conference on Human-Computer Interaction*, Springer, pp. 581–595 (2023).
- [14] Ward, J. A., Lukowicz, P., Troster, G. and Starner, T. E.: Activity recognition of assembly tasks using body-worn microphones and accelerometers, *IEEE transactions on pattern analysis and machine intelligence*, Vol. 28, No. 10, pp. 1553–1567 (2006).
- [15] Yan, H., Zhang, Y., Wang, Y. and Xu, K.: WiAct: A passive WiFi-based human activity recognition system, *IEEE Sensors Journal*, Vol. 20, No. 1, pp. 296–305 (2019).
- [16] Fu, B., Damer, N., Kirchbuchner, F. and Kuijper, A.: Sensing technology for human activity recognition: A comprehensive survey, *Ieee Access*, Vol. 8, pp. 83791–83820 (2020).
- [17] Ozcan, K. and Velipasalar, S.: Wearable camera-and accelerometer-based fall detection on portable devices, *IEEE Embedded Systems Letters*, Vol. 8, No. 1, pp. 6–9 (2015).
- [18] Tong, C., Ge, J. and Lane, N. D.: Zero-shot learning for imu-based activity recognition using video embeddings, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 5, No. 4, pp. 1–23 (2021).
- [19] Raju, V. and Sazonov, E.: Processing of egocentric camera images from a wearable food intake sensor, *2019 SoutheastCon*, IEEE, pp. 1–6 (2019).
- [20] Chen, J., Marshall, S. J., Wang, L., Godbole, S., Legge, A., Doherty, A., Kelly, P., Oliver, M., Patterson, R., Foster, C. et al.: Using the SenseCam as an objective tool for evaluating eating patterns, *Proceedings of the 4th international SenseCam & pervasive imaging conference*, pp. 34–41 (2013).
- [21] Nam, Y., Rho, S. and Lee, C.: Physical activity recognition using multiple sensors embedded in a wearable device, *ACM Transactions on Embedded Computing Systems (TECS)*, Vol. 12, No. 2, pp. 1–14 (2013).
- [22] Fernandes, G., Zhu, H., Pedram, M., Schauer, J., Shahi, S., Romano, C., Gergle, D. and Alshurafa, N.: Is cartoonized life-vlogging the key to increasing adoption of activity-oriented wearable camera systems?, *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–8 (2023).
- [23] Wang, J., Amos, B., Das, A., Pillai, P., Sadeh, N. and Satyanarayanan, M.: Enabling live video analytics with a scalable and privacy-aware framework, *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, Vol. 14, No. 3s, pp. 1–24 (2018).
- [24] Wang, J., Amos, B., Das, A., Pillai, P., Sadeh, N. and Satyanarayanan, M.: A scalable and privacy-aware IoT service for live video analytics, *Proceedings of the 8th ACM on Multimedia Systems Conference*, pp. 38–49 (2017).
- [25] Jha, S. S., Mayer, S. and Schneider, T.: Explicitly privacy-aware space usage analysis, *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, IEEE, pp. 1285–1290 (2020).
- [26] Chen, Y., Lai, Y.-K. and Liu, Y.-J.: Cartoongan: Generative adversarial networks for photo cartoonization, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 9465–9474 (2018).
- [27] Canny, J.: A computational approach to edge detection, *IEEE Transactions on pattern analysis and machine intelligence*, No. 6, pp. 679–698 (1986).
- [28] Li, B., He, F. and Zeng, X.: A novel privacy-preserving outsourcing computation scheme for Canny edge detection, *The Visual Computer*, Vol. 38, No. 12, pp. 4437–4455 (2022).