

令和 6 年 6 月 4 日現在

機関番号：82110

研究種目：若手研究

研究期間：2021～2023

課題番号：21K18109

研究課題名（和文）AIを利用した複数患者/入所者24時間アクシデント検知システムの開発と実用化

研究課題名（英文）Development of AI-based Multi-action Detection System for multiple patients using Day-Night Omnidirectional Cameras

研究代表者

今淵 貴志 (Imabuchi, Takashi)

国立研究開発法人日本原子力研究開発機構・福島研究開発部門 福島研究開発拠点 廃炉環境国際共同研究センター・研究職

研究者番号：90845471

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：本研究は、医療介護現場における患者や入所者のアクシデントとなり得る行動をAI（Artificial Intelligence）関連技術によって自動検知するモニタリングシステムについて研究開発を行った。全方位型カメラによって広範囲の映像を取得し、3次元関節座標抽出アルゴリズムを応用して算出した関節角度変動を用いて特定行動を推論するネットワークを学習した。提案手法群をプロトタイプシステムとしてエッジAI型コンピュータ上に実装して評価を行った結果、複数のアクシデント特定行動を高精度かつリアルタイムで検出可能であることを示した。

研究成果の学術的意義や社会的意義

全方位カメラで取得した2次元画像に対して人物の3次元関節座標抽出および関節変動角度の学習を組み合わせることで特定行動の高精度推定を実現した。本研究では、単眼の全方位カメラおよび小型エッジAIデバイスの組み合わせによりプロトタイプシステムを開発し、最小限のシステム構成でありながら広範囲をリアルタイムにモニタリング可能な性能を有していることを実証した。これは類似システムと比較して安価な設備投資で実用可能であり、医療介護現場のみならず他の分野においても応用・貢献できる。

研究成果の概要（英文）：In this study, we developed an automatic monitoring system for detecting potential accidental behaviors of patients and residents at medical care facilities using AI (Artificial Intelligence)-related technologies. Our system is characterized using an omnidirectional camera to acquire a wide range of video images and the application of a 3D joint coordinate extraction algorithm to predict specific accidental behaviors using calculated joint angle variations. In evaluation of prototype system that implemented the proposed methods on an edge-AI computer, we shown our system can detect multiple accidental behaviors with high accuracy and in real time.

研究分野：コンピュータビジョン

キーワード：Deep learning Human pose estimation Behavior analysis Omnidirectional camera

## 1. 研究開始当初の背景

超高齢化社会による医療介護サービスを必要とする患者の増加と医療介護従事者の不足の中、認知症患者や片麻痺患者の数は増加傾向にあり、当該患者の転倒・転落事故は大きな問題となっている [1]. 転倒や転落といったアクシデント行動の早期発見は重要であり、従来、24 時間体制の看護巡回が行われているが、医療従事者にとって大きな負担となるとともに、患者の眠りの妨げの要因となる。そこで、転倒や転落の行動を検知し通知することを目的に感圧センサや電磁タグセンサが医療介護現場へ導入されているが、これらは患者が意図的に避ける・外すといった欠点がある [2]. また、これらは患者に”拘束”の不快感を与えることに繋がる。一方、ICT (Information and Communication Technology) 技術の導入により、監視カメラを応用したモニタリングシステムは普及してきており、市販サービスも存在が、匿名性の確保が難しいという問題がある [3]. 深度カメラセンサによる自動非接触型システムもあるが、モニタリング範囲が狭く、広範囲をカバーするために設備投資が大きくなる課題がある [4]. 以上のことから、モニタリングシステムは、設置が容易で広範囲を一度に計測可能なハードウェアを用いて、24 時間をとおして人物の特定行動を検知し、匿名性を確保した単純なデータを利用するものが望ましいと言える。患者の心理的および身体的側面を抑制しない新しい光学モニタリング装置の実用化が必要とされている。

## 2. 研究の目的

本研究では、昼夜を通して、同一空間内の複数患者のアクシデントとなり得る行動を自動検知するモニタリングシステムの実用化に向けて、行動検知アルゴリズムおよびシステム構成について研究開発を実施する。昨今の AI 関連技術における 2 次元関節座標推定アルゴリズムは、カメラから 5m 以上離れた対象人物の関節座標を追跡する能力を有し、多くの分野で活用が進んでいる。これまで我々は、ステレオ全方位カメラを用いて当該 AI 関連技術を応用した 3 次元関節座標抽出に基づいたルールベースの行動検知システムを試作しているが、行動検知の精度、頑健性とシステム導入コストに課題がある [5]. そこで、単眼の全方位カメラを利用することで全体のコスト削減を図る。これに伴い、行動検知アルゴリズムにおいては 3 次元関節座標推定アルゴリズムを導入し、転倒や徘徊といった行動について関節情報を時系列に学習する深層学習モデルの導入を検討する。さらに、エッジ AI 型デバイスを用いたプロトタイプシステムの開発によって実用性を評価することを目的とする。

## 3. 研究の方法

カメラセンサで取得した動画像から人物の行動を推定する場合、アピランスに基づき画像特徴量を直接的に学習する手法が存在するが、背景オブジェクトや暗所映像下におけるノイズの影響で誤推論が起きやすい。さらに、全方位カメラのように広範囲を一度に取得して展開した画像では、通常のカメラ画像と比較して画素数が低く、かつ、歪みが生じるため、撮影した人物が遠距離になるほど検知精度が低下する。そこで本研究では、カメラ画像から抽出した人物の関節情報を用いて行動検知を行うアプローチをとる。システムの入力は、全方位カメラで取得した画像を 4 分割射影変換により展開したパノラマ画像とする。はじめに、人物の骨格情報を学習済のネットワークおよびカメラパラメータを用いてパノラマ画像から 3 次元関節座標を抽出、さらに関節座標から関節角度を算出する手法について研究開発を行う。次に、複数の関節角度の時系列の変動をアクシデント行動毎に深層学習モデルによって学習することで、アクシデント行動の検知を行う手法について研究開発を行う。3 次元の関節座標を直接学習する場合、各行動を表現する関節の自由度が高く、また、背景オブジェクトやノイズによって座標が大きく変動してしまうことで高精度での学習が難しい。ここでは、関節座標から関節角度の変動を計算して学習することで、人物の向きやカメラ画像位置に依存せず高精度な行動の学習と検知を可能とする。これらの開発手法群に基づいて、エッジ AI 型デバイスを用いたプロトタイプシステムを開発する。さらに、アクシデントが検知された際に、関節情報や検知結果、時間や場所のメタデータとパッケージ化して医療従事者が持つ端末への情報送信する機能についても実装を行う。これは、デバイス上で動画像を用いた計算処理を完結させることで、実利用においては匿名性を確保できるためである。モニタリングシステムの実用の観点からは、行動推論の精度および処理のリアルタイム性が重要となるため、本研究ではこの 2 点について評価実験を行う。

## 4. 研究成果

### (1) 全方位型アクシデントモニタリングシステムのための行動検知手法

上述の研究の方法に基づいて、全方位カメラによるアクシデントモニタリングシステムの行動検知手法を提案し、評価実験を行った結果、6 行動を高精度に検知可能であることを示した。

図1は、提案手法の概要を示す。まず、双曲面ミラーレンズによる全方位カメラで取得した動画像を4分割射影変換によって4組の画像へと展開して、パノラマ画像を作成した。ここで、全方位カメラは事前にカメラキャリブレーションを行っており、内部パラメータは既知とする。次に、パノラマ画像から人物の骨格情報を抽出する。本研究では、Mediapipe [5]のPose moduleを使用して胴体、腕、脚、顔にまたがる17個の3次元関節座標を取得した。Mediapipeは、Googleが開発したAIライブラリであり、カメラの内部パラメータと2次元関節座標から学習済みモデルを用いて3次元関節座標を推定する。続いて、3次元関節座標に基づいて、8関節（右肩、左肩、右肘、左肘、右股関節、左股関節、右膝、左膝）について、それぞれの支点となる3次元関節位置座標から隣接する関節座標を構成する2つの3次元関節ベクトルを用いて関節角度を算出した。関節角度の変動は時系列に記録され、行動検知ネットワークへ入力される。行動検知ネットワークは、エッジAI型デバイスでの動作を想定して、計算コストが低く、時系列の学習が可能なLSTM (Long-Short Term Memory) モデルを採用して行動の学習を行った。図2に行動検知ネットワークの概要を示す。ネットワーク構成は、8関節の過去20フレームの関節角度の変動を入力として、20ユニットのLSTM層×2、Dense層、Softmax層で構成される。出力層は6行動（立位、座位、臥位、立上、転倒、徘徊=歩行）の推論確立を出力することとした。また、損失関数はCross-Entropy Errorとした。行動検知ネットワークの検出結果と予めパノラマ画像上で設定したベッド領域、ドア位置等を利用することでベッドからの転倒等のアクシデントを設定することができる。

行動検知ネットワークの学習は、1人の被験者により6行動について各50回の動作を全方位カメラで撮影した動画像（2448×2050画素、10fps）を用いて、正解ラベルを付与した教師データセットを作成して学習を行った。図3は動画像の撮影環境および各行動における人物領域の画像と関節角度変動の一例を示す。各行動50回分のデータを収集して、データ拡張として映像を左右反転させたデータを加えた合計600データを学習および検証に用いた。学習設定は、エポック数を100、バッチサイズを14、最適化アルゴリズムをAdaptive moment estimation (Adam)として、モーメント0.1、初期学習率は $1e-4$ とした。作成したデータセットのうち70%（420データ）を学習、残り30%（180データ）をテストに用いた。学習に用いた計算機は、CPU: Intel Core i9-11900K 3.5GHz, GPU: Nvidia GeForce RTX3090 24GB, memory: 64GBの仕様であった。テストデータに対する推論の結果、精度93.3%、適合率94.1%、再現率93.3%およびF1値93.1%であった。図4はテストセットにおける検知精度の混同行列を示す。「立位」および「立上」を除いて100%と高精度であることを確認した。「立位」については、一部テストデータが「歩行」および「臥位」に誤検知される結果となったが、前者は歩行時の肢体の振り幅が少ないことから関節角度が類似したこと、後者も同様に関節角度の類似が要因であると考えられる。今回の評価は、特定の被験者に対する学習を想定したものであり、高精度の行動検知を実現したが、実利用に向けては、大規模なデータセットを用いたロバストな識別器を学習することや、個人の日常行動から行動特性を算出して識別器をチューニングする方法等の工夫が必要といえる。今後、これらの方法の研究開発および検知モジュールの頑健性向上のために最新の深層学習バックボーンを用いた性能比較を実施していきたい。

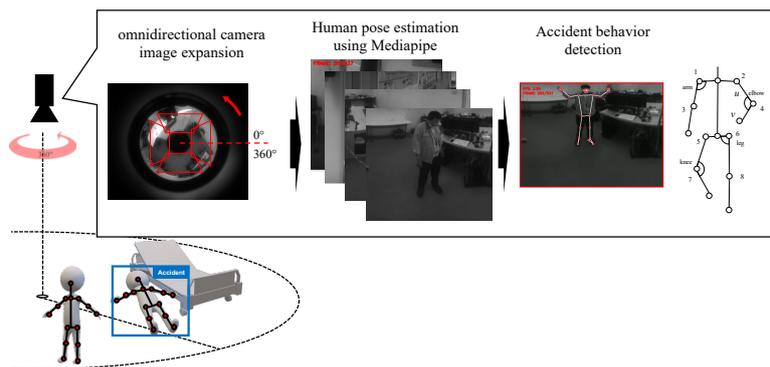


図1 全方位カメラによるモニタリングシステムの概要

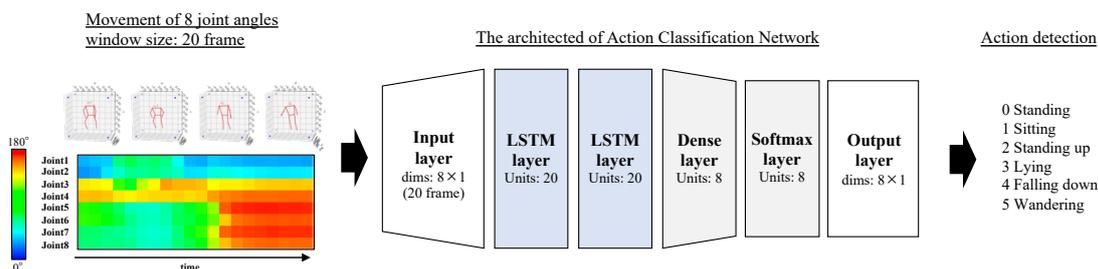


図2 行動検知ネットワークの概要

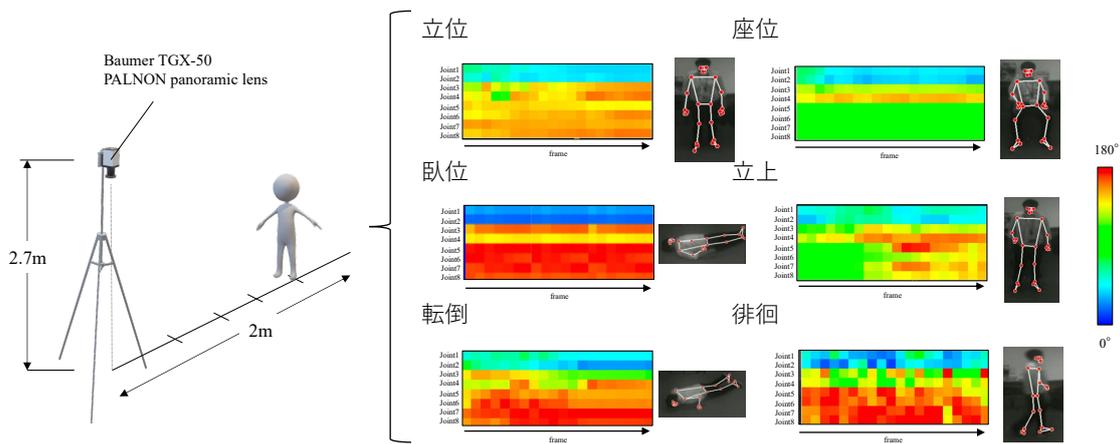


図3 性能評価における撮影環境と6行動における関節角度変動の一例

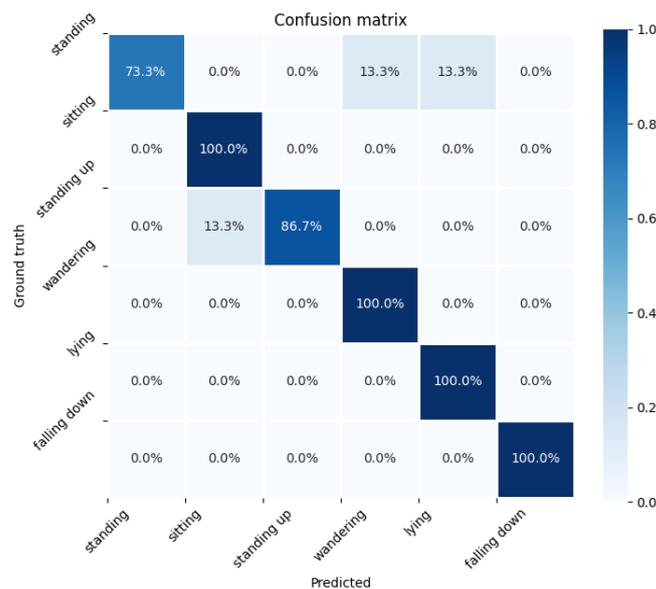


図4 テストセットにおける検知精度の混合行列

## (2) エッジAI型デバイスによるプロトタイプシステムの開発

(1)の行動検知手法をエッジAI型デバイス上に実装してプロトタイプシステムを開発し、性能評価を行った結果、リアルタイムで動作可能であることを示した。

図5はプロトタイプシステムを天井付近に設置した様子を示す。モニタリングシステム全体の設備導入コストを低くするために、市販のエッジAI型デバイスと単眼の全方位カメラの使用、さらにPoEによる給電を想定する。本研究では、研究開発用デバイスとして、Baumer社製のVAX-50C. I. NVX [7]上に手法群を実装した。当該デバイスは、NVIDIA Jetson NanoおよびCマウントの工業用カメラマウントを備えたデバイスであり、CUDAコアを用いた深層学習モデルの動作が可能である。また、Gigabitイーサネットの装備、ヒートシンクによる耐熱構造のためエッジデバイスとして連続稼働によるモニタリングとデータ転送が可能である。当該デバイスに双曲面ミラー式全方位レンズを装着して画像を取得した。全方位画像の解像度は $2048 \times 1536$ であり、最大55fpsで取得可能である。OSは、Jetpack 4.5でありソフトウェアはPythonによって実装した。パノラマ画像の展開および行動識別器は、それぞれOpenCV. 4.0およびTorch 1.8の各種メソッドを用いて実装を行なった。ここで、全方位カメラ画像の展開の計算コストを減らすため、4分割画像の射影画像から歪み係数および展開中心座標を用いて同心円上に画像変換を行う展開方式に変更を行った。検知した人物の位置座標やアクシデント行動はネットワークを通して逐次送信する仕様とした。プロトタイプシステムの動作検証として、ローカルネットワークを介して行動検知結果をタブレット端末へ通知して表示するためのビューワについても開発を行った。

ここで、プロトタイプシステムの行動検知性能を明らかにするために、(1)で述べた実験と同条件下で実験を行った。行動検知ネットワークはPyTorchによって作成し、Nvidia GeForce

RTX3090 を搭載したデスクトップ端末により学習を行い、テストにおいて当該学習済モデルを用いてエッジ AI 型デバイス上で推論を実施した。その結果、アクシデントとなり得る行動である転倒、臥位、徘徊の検知率はそれぞれ 100%、100%および 93.3%であり、6 行動の平均精度は、97.8%であった。誤推論傾向は(1)の実験と同様の傾向が見られたが、エッジ型 AI デバイス上でも同等な推論が可能であることを確認した。図 6 は、検知結果をタブレット端末上で表示した結果の一例である。ここでは、実験結果の確認用にカメラ映像を送信して映しているが、実利用の際は通知内容を行動検知結果や時間、場所の提示のみに限定することで匿名性の確保が可能になる。エッジ AI 型 デバイス上に実装した各モジュールの処理全体を統合した結果、処理速度は約 3 fps であった。当該処理速度は十分とは言えず、特定行動を取り逃がす場合が生じる懸念がある。今後、さらなる実用性の向上のために、処理コストが高い全方位画像のパノラマ展開と行動検知について、パフォーマンスの改善が必要である。また、今回は事前に撮影した動画像を用いた評価実験であったが、カメラセンサからの直接入力を用いた場合に同等な結果を得ることができるか確認が必要である。設備導入コストについては、S マウントレンズおよび高解像度 Web カメラの組み合わせを採用することでコストを大きく抑えることが期待できる。これについても、本研究と同等な結果が得られるかについて今後評価を行う必要がある。



図 5 プロトタイプシステムの外観



図 6 検知結果表示の一例

#### <引用文献>

- [1] I. D. Cameron, S.M. Dyer, C.E. Panagoda, “Interventions for preventing falls in older people in care facilities and hospitals,” Cochrane Database Syst Rev, 2018.
- [2] I. Kitayama, K. Omori, H. Matsuno, and Y. Sugimoto, “Wandering prevision and monitoring systems for persons with dementia (Part 2) - Usage survey of wandering prevision and monitoring system/machine at nursing homes and other welfare facilities-,” Report of the Hyogo Prefectural Town Welfare Research Institute, pp. 98-111, 2003.
- [3] S. T. L. Pohlmann, E. F. Harkness, C. J. Taylor, and S.M. Astley, “Evaluation of Kinect 3D sensor for healthcare imaging,” Journal of Medical and Biological Engineering, Vol.36, pp. 857-870, 2016.
- [4] 介護施設向けサービス HitomeQ ケアサポート, <https://www.konicaminolta.com/jp-ja/caresupport/index.html>
- [5] T. Imabuchi, Y. Murata, and Prima O.D.A, “Multiaction Detection System Using Infrared Omnidirectional Cameras,” The Twelfth International Conference on eHealth, Telemedicine, and Social Medicine (eTELEMED 2020), 2020.
- [6] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C. Chang, M. Yong, J. Lee, W. Chang, W. Hua, M. Georg, M. Grundmann, “MediaPipe: A Framework for Perceiving and Processing Reality,” Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR), 2019.
- [7] Baumer, VAX-50C. I. NVX, <https://www.baumer.com/us/en/product-overview/industrial-cameras-imageprocessing/industrialcameras/ax-series/nvidia-jetsonxavier-nx/vax-50c-i-nvx/p/44269>

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 今淵 貴志、プリマ・オキ ディッキ・アルディアンシャー
2. 発表標題 全方位型アクシデントモニタリングシステムのための深層学習を用いた行動検知
3. 学会等名 第50回画像電子学会年次大会
4. 発表年 2022年

1. 発表者名 今淵 貴志
2. 発表標題 全方位カメラを用いたエッジAI型アクシデントモニタリングシステムの開発
3. 学会等名 情報処理学会第86回全国大会
4. 発表年 2023年

〔図書〕 計0件

〔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
---------------------------	-----------------------	----

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
---------	---------