

## 骨格検出を用いた行動認識技術

根尾敦 小林伊織 大宮英典

株式会社日立製作所

## 1. はじめに

センサの飛躍的な能力向上、および数量の増加に伴い、データ処理における通信量や遅延量等の低減を目的として、エッジコンピューティングが着目されている[1]。特にセキュリティ需要の高まりから防犯カメラの導入実績は多く、例えば東日本旅客鉄道（JR 東）は、在来線の約8,300両と、新幹線200両に防犯カメラを2018年夏以降、順次追加設置してきた[2]。しかし設置された監視カメラの数は多く、労働人口の減少が問題視される中[3]、防犯カメラの映像監視は人手作業ではなく、エッジによる画像認識で実施されることが望ましい。本報告では、エッジでも動作可能な処理負荷の低い、人の行動認識技術の検討および評価を行う。

## 2. 行動認識技術の方式選定

本報告ではエッジ向けに処理負荷の低い行動認識技術を検討する。このため、様々な組織で開発が進められ、エッジ向けに処理負荷の低い手法が発表されている骨格検出技術に着目した[4]。着目した骨格検出技術では18点の人の骨格位置を $(x, y)$ の2次元座標形式で画像から推論する。人の行動は各骨格の互いの位置関係に密接な関係があると考え、検出した骨格位置から人の行動を推論する。

しかし、18点の2次元座標を適切に行動に結びつけるには変数が多いことから処理負荷が高いことと、処理の煩雑さが予想される。このため、本検討では主成分分析[5]による次元縮約と、決定木[6]による行動分類に着目した。

## 3. 行動認識アルゴリズム

主成分分析とは元のデータ $x_i$ から次元の縮約により新しい変数 $P$ を生成して、元のデータの特徴を分析する手法である(式(1))。

$$P_1 = w_1x_1 + w_2x_2 \cdots + w_px_p \quad (1)$$

( $w_i$ は次元縮約のための係数)

例えば主成分分析により2個の変数を生成した場合、図1に示す通り、行動を平面上に配置できる。本行動認識アルゴリズムでは、主成分分析後の変数空間上で、決定木により行動分類を行うことで行動認識を行った。

検討した行動認識アルゴリズムを図2に示す。本行動認識アルゴリズムでは骨格検出と行動分類で事前学習が必要になる。骨格検出では画像と画像に映る骨格座標が、行動分類では骨格座標と行動ラベル情報が教師データとして必要になる。行動ラベル情報は教師データとして用いる骨格情報がどういった行動を示すものかを示す情報である。

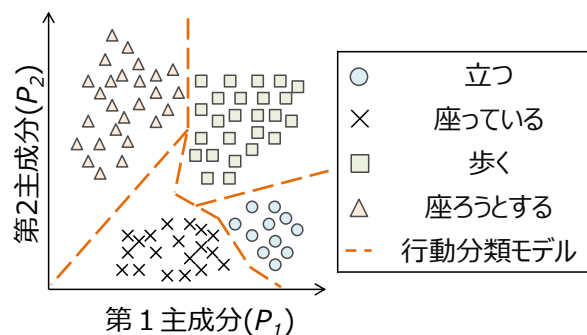


図1 主成分分析後の行動分類

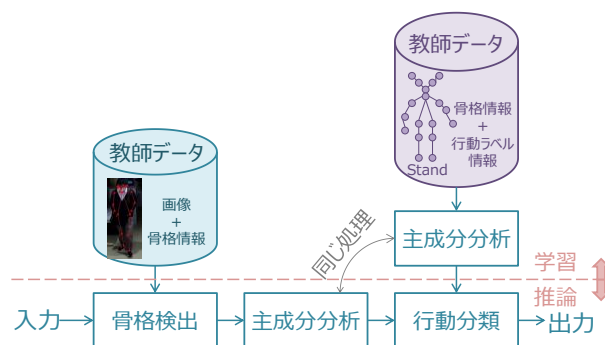
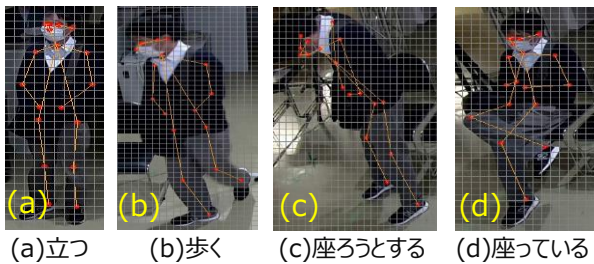


図2 行動認識アルゴリズム

## 4. 評価

認識精度評価は図3に示す4種類の行動を対象に、独自に用意した360枚の画像で行った。また処理速度の評価として、図4に示す日立制御エッジコンピュータ CE50を用いた。



(a)立つ (b)歩く (c)座ろうとする (d)座っている

図3 評価対象の行動

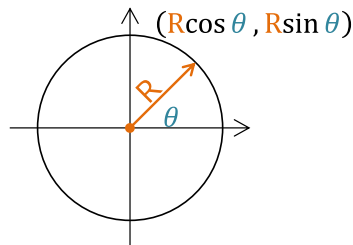


Intel Atom®  
x7-E3950  
(1.6GHz/4Core)搭載

図4 日立制御エッジコンピュータ CE50



格子間隔は  
10Pixel



R	引数で指定
θ	乱数発生 $0 \leq \theta < 2\pi$

18点各点にノイズ印加

図5 評価時のランダムノイズ

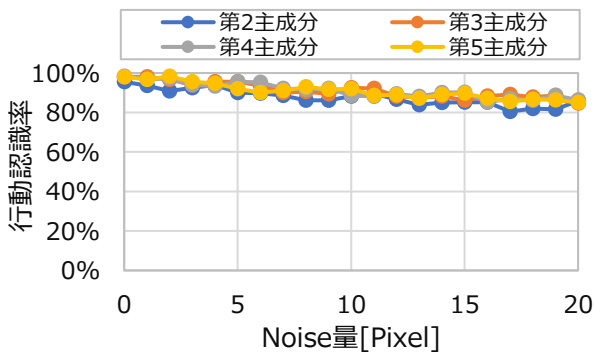


図6 4行動分類結果

表1 処理速度

デバイス	CPU	GPU
速度 [fps]	2.6	10.5
CPU 使用 [%]	91.0	34.9
Memory [GB]	2.4	2.8

行動認識精度は骨格検出の精度に左右されるため、本検討では教師データに用いた骨格座標にランダムノイズを付加することで骨格検出精度も考慮した行動認識精度を評価した。ランダムノイズの印加方法を図5に示す。ノイズ量Rに乱数生成器により取得した角度 $\theta$ を組み合わせ、骨格18点の各点にノイズを付加した。

行動認識精度を図6に示す。凡例は主成分分析後の変数をいくつ用いたかを示し、第3主成分は第1主成分から第3主成分までの3つの変数を行動分類に用いたことを示す。骨格毎に20ピクセルのノイズを与えても80%以上の精度で行動認識可能である。

処理負荷について処理速度とCPU使用率で評価した(表1)。評価に用いたIntel Atom® E3950はファンレスで実装可能な組込向けのCPUである。このため一般的なPCに用いられるCPUより処理能力が低く、骨格検出を含む全ての行動認識処理をCPUに行わせた場合、CPU使用率が高い。そこで本検討では骨格検出をGPUにオフロードすることで、CPU負荷を下げ、処理速度を向上させた。

## 5. まとめ

本稿では、エッジ向けの行動認識技術として、骨格検出と主成分分析と決定木を用いるアルゴリズムを考案し、4行動を対象に認識精度と処理負荷を評価した。その結果ノイズ付加の状況においても低負荷で80%以上の高精度な精度結果を得た。今後は、認識対象とする行動の多様化が課題になる。

## 参考文献

- [1] 山口弘純, "IoT エッジセンサの知能化に向けて", 分散協調とモバイルシンポジウム 2017 論文集 (2017)
- [2] "JR 東日本ニュース", 2018-07-03
- [3] 西村俊, "パーソナルツールズをターゲットとした地域資源の再構成", 民族植物学ノオト (13), 10-16, 2020-03-30
- [4] Zhe Cao, "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields", CVPR 2017
- [5] Pearson, K.. "On Lines and Planes of Closest Fit to Systems of Points in Space", Philosophical Magazine 2 (11): 559-572, 1901
- [6] L. Breiman, "Classification and Regression Trees", Wadsworth (1984)