

歩き方の特徴に着目した所持物の種類・重さ認識の検討

水野 雅也^{1,a)} 川西 康友^{1,2,b)} 出口 大輔^{1,c)} 村瀬 洋^{1,d)} 井下 哲夫^{1,e)}

概要

重い荷物を持っている人物は、軽い荷物を持っている時とは歩き方が異なる。本発表では、荷物を持った人物の歩き方に着目することで、その人が所持している荷物の種類と重さを認識する手法について検討する。提案手法では、この課題に対して時空間グラフ畳み込みネットワーク (ST-GCN) を応用する。しかし、ST-GCN は歩行速度の違いに起因する系列長の違いをうまく扱えない。この問題に対して、ランダムサンプリングによって得た複数の部分系列に対する認識結果を統合することで、可変長の系列を扱う手法を提案する。実験の結果、荷物の種類と重さを同時に認識する 10 クラス分類が 71.1% で実現できた。

1. はじめに

スポーツ会場や地下鉄など大人数が集まる施設にはテロの危険性がある。このような施設でのテロを未然に防ぐ方法の 1 つとして、施設を出入りする人物に対して不審物の持ち込みを取り締まる事が挙げられる。不審物の持ち込みを取り締まる場合、その施設を利用するすべての人が通る出入口で取り締まりを行なうことが良いと考えられる。そこで本研究では、出入口の監視カメラを用いて不審物の持ち込みを取り締まることを考える。

ここで、テロに用いられる不審物には銃などの武器や爆弾、燃料など、重いものが多いと考えられる。また、不審物はリュックサックや手さげかばんなどの所持物に隠して持ち運ぶと考えられる。そのため、所持物の種類・重さを認識することができれば不審物の検出に貢献できる。本研究では所持物の種類・重さを認識することを目的とする。

所持物認識の既存手法として、人物画像から所持物を持っているか否かを判定する手法 [1]、観測したシルエットと、人物骨格から推定したシルエットとの差を用いて所

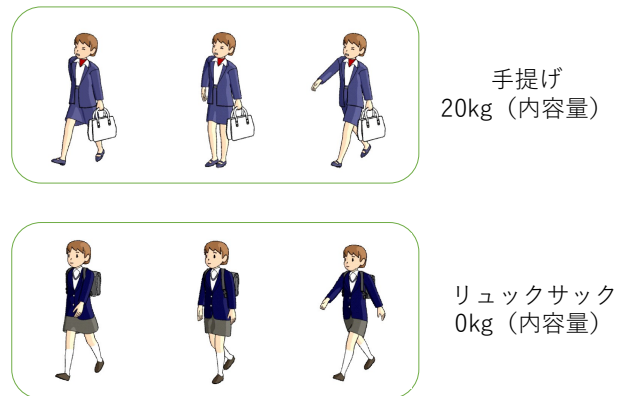


図 1 歩き方から所持物・重さを認識する概念図

持物を持っているか否かを判定する手法 [2]、人物骨格の手の位置から所持物の存在領域を設定し、物体検出結果を絞り込むことで誤検出を減らす手法 [3] がある。しかしこれらの研究では、所持物自体が遮蔽される場合には検出できないという問題点がある。また、重さの違いの認識は対象外である。

一方、所持物の重さを認識する研究として、身体動揺を用いて所持物の重さを認識する研究 [4] がある。身体動揺とは、人物が静止している時に生じるわずかな体の揺れである。この研究では、所持物が重いほど身体動揺が大きいことを利用して所持物の重さ認識を行なう。しかし、身体動揺は静止している人を頭上真上から撮影する必要があるため、今回想定する状況では難しい。

これらの問題に対し本研究では、重いものを持って歩くときと軽いものを持って歩くときでは歩き方が異なるという点に着目する。そして、歩き方の特徴を用いて所持物の種類・重さを認識する手法を提案する。

歩き方の特徴として、歩行は 2 歩分 (歩行 1 周期分) の繰り返しであることに着目し、歩行 1 周期分の人物骨格系列を利用する。歩行 1 周期分の人物骨格系列は、背景の影響を受けない、想定する環境において遮蔽されづらいという利点がある。

本研究では歩き方の特徴として歩行 1 周期分の人物骨格系列を用いるが、(i) 歩行 1 周期分の人物骨格系列はそれぞれ長さが異なる可変長系列であるという問題点がある。そのため、系列同士を直接比較することが容易でない。また、

¹ 名古屋大学

² 理化学研究所 情報統合本部 GRP

³ 日本電気株式会社

a) mizunom@vialab.is.i.nagoya-u.ac.jp

b) yasutomo.kawanishi@riken.jp

c) ddeguchi@nagoya-u.jp

d) murase@nagoya-u.jp

e) tetsuo.inoshita@nec.com

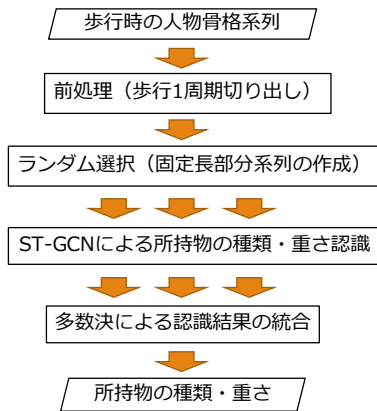


図 2 提案手法の処理手順

(ii) 関節座標の集合による人物骨格系列のみでは、各関節の隣接関係や時間方向の変化の表現が不十分であるという問題点もある。これらの問題に対し、本研究では RTW [5] の考え方を応用し、歩行 1 周期分の系列からフレームを一定数ランダムに抜き出すことで固定長部分系列を作成するランダム選択を導入する。これにより歩行 1 周期の長さの違いを吸収することで、(i) の問題に対処する。また、各固定長部分系列の認識のために ST-GCN [6] を導入することで、人物骨格系列を時空間グラフにより表現し、グラフ畳み込みを行なう。これにより人物骨格構造と姿勢の時間変化を考慮した特徴抽出を実現し、(ii) の問題に対処する。

2. 歩き方の特徴を用いた所持物の種類・重さ認識

2.1 提案手法の概要

本項では、歩行 1 周期分の人物骨格系列によって表現された人物の歩き方から、所持物の種類・重さに対応するクラスに分類を行なう手法について述べる。図 2 に提案手法の処理手順を示す。

以下で本手法のポイントである、ランダム選択と、人物骨格の構造を考慮した ST-GCN を用いた認識処理について詳述する。なお本手法では、斜め上から撮影された人物の歩行時の人物骨格系列 X を入力とする。前処理として歩行時の人物骨格系列 X から歩行 1 周期分 $X_i = \{x_1, x_2, \dots, x_{N_i}\}$ を切り出しておく。

2.2 ランダム選択

提案手法では、歩行 1 周期分の人物骨格系列に対してランダム選択を複数回行なうことで固定長部分系列を複数作成する。このとき、選択したフレームの時間順序は変えない。ランダムに選択することにより、時間的な揺らぎの多様性を持たせる効果もある。また、複数回行なうことで、単純に等間隔で切り出して固定長系列を作成するのに比べ、万遍なく時系列の情報を活用できる。

ランダム選択の処理の例を図 3 に示す。歩行 1 周期分の

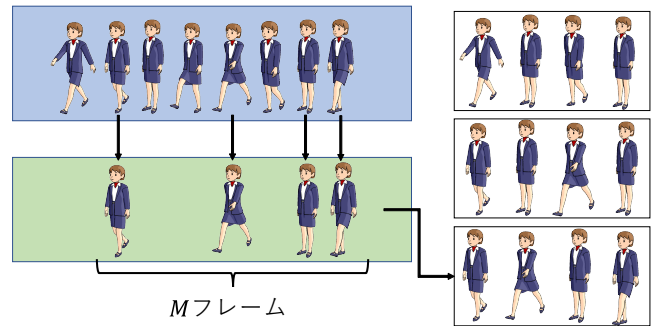


図 3 ランダム選択の例

系列 X_i から重複なくランダムに固定長 $M (M < N_i, \forall_i)$ 分のフレームを抜き出し部分系列 $X_{ij} (X_{ij} \subset X_i, |X_{ij}| = M)$ とする。この処理を P 回繰り返すことで、 X_i から P 個の固定長部分系列 $X_{ij} (j \in \{1, \dots, P\})$ を得る。

2.3 ST-GCN の出力の多数決による認識

提案手法では、ランダム選択で作成した個々の固定長部分系列に対して所持物の種類・重さ認識を行なう。ST-GCN による分類を行なうことで、人物骨格系列の骨格構造と時間変化を考慮した特徴抽出を実現する。まず、ランダム選択により得られた P 個の固定長部分系列 X_{ij} に対し、同一フレーム内の人体の構造的に接続している関節である隣接関節と前後フレームの同一関節を接続することで無向グラフ G_{ij} を得る。次に、 G_{ij} をそれぞれ所持物の種類・重さ認識器 f に入力する。そして、所持物の種類・重さ認識結果 $a_{ij} = f(G_{ij}; \theta)$ を P 個得る。ここで、 θ は f のパラメータである。最後に、ST-GCN による認識により得られた P 個の認識結果 a_{ij} を多数決 V により 1 つの認識結果 $a_i = V(\{a_{ij} | \forall_i\})$ に統合する。

本実験で使用する所持物の種類・重さ認識器 f の構造を図 4 に示す。グラフの下の数字はチャンネル数を表す。ST-GCN と全結合層の活性化関数には Leaky ReLU を用いる。出力層の活性化関数には Log Softmax を用いる。入力は G_{ij} であり、所持物の種類・重さ認識を $f(G_{ij}; \theta)$ と表現する。 θ は学習により求める。

3. 実験

3.1 データセット

歩行時の人物の様子を撮影し、その人物の所持物の種類・重さを真値として付与した公開データセットはないため、本研究では実験用データセットの撮影を行なった。本節ではデータセットを構築した際の撮影条件及びその内容について述べる。

本研究では、1 人の歩行者が 1 つの所持物を持って歩いている様子を歩行者の前方から見下ろすように観測する状況を想定する (図 5)。実験参加者に荷物を所持して歩かせ、その様子を 2 m の高さに設置した Microsoft 社製の

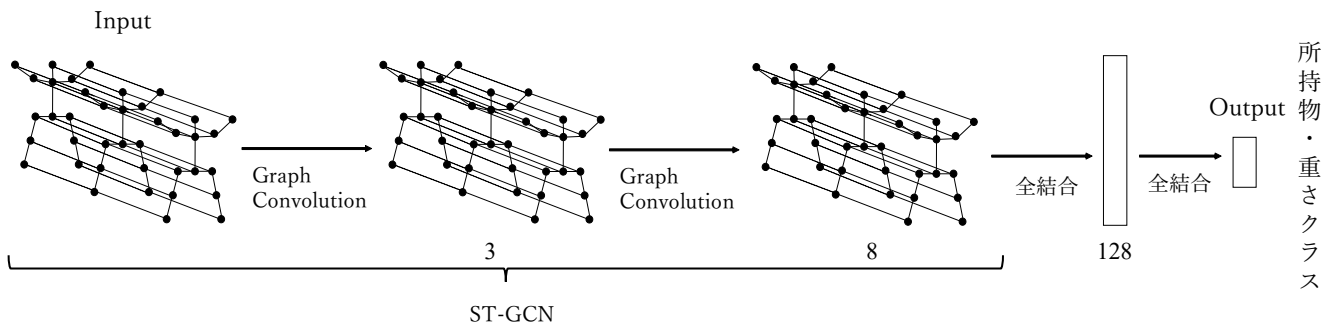


図 4 部分系列に対する所持物の種類・重さ認識器の構造

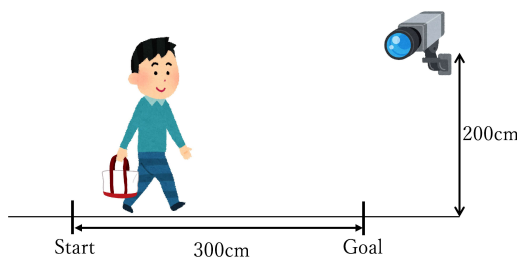


図 5 撮影環境の断面図

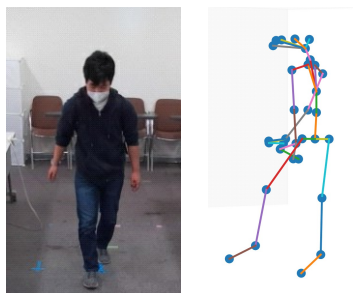


図 6 撮影画像と人物骨格の例

Azure Kinect DK で撮影することで 3 次元の人物骨格系列を取得した。撮影した画像と人物骨格の例を図 6 に示す。人物骨格は、カメラ座標系における各関節点の (X, Y, Z) 座標値の集合である。本研究では、フレームごとに腰の関節点座標を $(0, 0, 0)$ とし、背中の中長さが 1 になるよう正規化を施した。実験参加者は 9 名（男性 7 名、女性 2 名）であり、所持物の種類は「手ぶら」、「手さげかばん」、「肩さげかばん」、「リュックサック」とし、内容を「0 kg」、「10 kg」、「20 kg」とした。また、全実験参加者は撮影データの研究目的での利用・公開に同意した。

3.2 実験手順

本実験では重さは認識せず、所持物の種類のみを認識する 4 クラス認識と、重さと所持物を同時に認識する 10 クラス認識、種類を既知として各所持物ごとに重さを認識する 3 クラス認識の 3 種類の認識を評価する。評価実験の評価指標として、認識率に対して正解率 $Acc = c_c / c_T$ を用いる。なお、 c_T は評価用データに含まれる歩行 1 周期の数を

表 1 比較した手法

	ランダム選択	人物骨格系列の構造
ST-GCN	×	○
RTW+NN	○	×
RTW+ST-GCN (提案)	○	○

表し、 c_c はそのうち正しく認識できた数を表す。本実験では 9 人の被験者のうち、1 人を評価用、残りを学習用を使用する交差検証である Leave-One-Person-Out によって精度を求めた。

提案手法におけるランダム選択と人物骨格系列の構造のそれぞれの効果を確認するため、提案手法を含め表 1 に示す 3 手法を比較した。ランダム選択では、抜き出すフレーム数は認識精度が高かったことから 10 フレームとし、固定長部分系列を 60 個作成した。ランダム選択を用いない場合は、歩行 1 周期全体から等間隔に 20 フレームを抜き出して固定長系列とした。ST-GCN を用いない場合は、全結合ニューラルネットワークを使用した。

3.3 実験結果

表 2 に所持物の認識と所持物・重さ認識の正解率を、表 3 に所持物ごとの重さ認識の正解率を示す。実験結果から、提案手法を用いることで、リュックサックの重さ認識以外のすべての認識において、正解率が比較手法より優れていることが分かる。

3.4 考察

表 2 に示した所持物の認識結果から、所持物の種類は視認できずとも人物の姿勢から認識できることが分かった。また所持物・重さの認識結果と表 3 から、所持物の重さも人物の姿勢から認識できることが分かった。一方で、図 7 を見ると、同じ種類・重さの所持物の場合でも人により姿勢に差があることが分かる。また、人物 A が 20 kg のリュックサックを持つときと人物 B が 10 kg のリュックサックを

表 2 所持物のみと所持物・重さ認識の結果

		所持物のみ	所持物・重さ
比較手法	ST-GCN	85.4%	54.4%
	RTW+NN	88.6%	68.2%
提案手法	RTW+ST-GCN	89.0%	71.1%

表 3 各所持物ごとの重さ認識の結果

		手提げかばん	肩さげかばん	リュックサック
比較手法	ST-GCN	63.6%	69.0%	60.5%
	RTW+NN	78.7%	80.6%	68.1%
提案手法	RTW+ST-GCN	78.8%	86.7%	66.1%

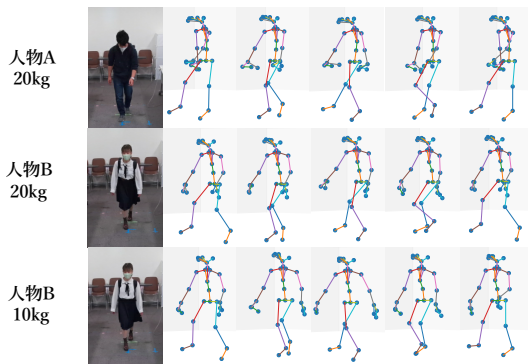


図 7 リュックサックを所持物とする場合の人物骨格例

持つときの姿勢が似ていることから、種類が同じで重さが異なる所持物の場合に人物の違いによって姿勢が似ることがあると分かる。これは、人物ごとに体格や筋肉量が異なるためであると考えられる。提案手法では認識に人物骨格系列のみを使用しており、人物の体格や筋力量を表現できていないため、今後は人物の体格や筋力量を考慮した特徴を含めた認識手法を提案する必要があると考えられる。

また、本実験では Azure Kinect DK を用いて人物骨格を推定したが、骨格推定を誤ったフレームが存在した。これは、被験者がカメラから 2m 以上離れてしまう場合や、被験者がスカートなど体の線を隠してしまう服装である場合に多く発生していることが分かった。本実験ではこれを人手で除去したが、実用の際は過去フレームとの比較や対象者の位置などを用いて自動的に除去できるようにする必要がある。また、対象者の服装によって骨格推定が失敗してしまう問題を解決するために、服装の違いに頑健な骨格推定器を開発する必要がある。

4. むすび

既存研究では歩行者の所持物の種類を認識するものがほとんどであり、その重さを認識していない。そこで本研

究では歩き方の特徴から、所持物の種類・重さを認識する手法を提案した。歩行 1 周期はそれぞれ長さが異なり、人物骨格系列は隣接関係や時間方向のつながりが考慮されていないため、ランダム選択により歩行 1 周期から複数の固定長部分系列を作成し、ST-GCN に基づく認識器により、所持物の種類・重さ認識を行なった。

今後の課題として、個人差への対応や人物骨格推定精度の向上、人混みへの対応などが挙げられる。

謝辞 本研究の一部は JSPS 科研費 (17H00745) による。

参考文献

- [1] 浅井康博, 西堀研人, 高橋友和, 出口大輔, 井手一郎, 村瀬洋, 複数方向から撮影された人物の手荷物所持判定に関する検討, 信学技報 (PRMU), vol.114, pp.21-24, Jan 2015.
- [2] 三浦誠, 阿部亨, 菅沼拓夫, 人物の骨格情報を用いた所持物検出手法に関する一検討, 情報処理学会第 80 回全国大会講演論文集, pp.445-446, Mar 2018.
- [3] 西田尚樹, 川西康友, 出口大輔, 井手一郎, 村瀬洋, 谷内田尚司, 白状の性質に着目した白状利用者検出に向けた検討, 2018 年電子情報通信学会総合大会講演論文集, pp.76, Mar 2018.
- [4] 山口優太, 神谷卓也, 西山正志, 岩井儀雄, 身体動揺を用いた重量物所持の認識可能性の検証, ビジョン技術の実利用ワークショップ講演論文集, no.IS2-A1, pp.1-5, Dec 2019.
- [5] C. H. Suryanto, J-H. Xue, K. Fukui, Randomized time warping for motion recognition, Image and Vision Computing, vol.54, pp.1-11, Oct 2016.
- [6] S. Yan, Y. Xiong, and D. Lin, Spatial temporal graph convolutional networks for skeleton based action recognition, Proc. of the 32nd AAAI Conf. on Artificial Intelligence, pp.7444-7452, Feb 2018.