

# 複数投影視点の適応的学習に基づく 3D LiDAR を用いた歩容認証

安正鎬<sup>1,a)</sup> 中嶋 一斗<sup>1,b)</sup> 吉野 弘毅<sup>1,c)</sup> 岩下 友美<sup>2,d)</sup> 倉爪 亮<sup>1,e)</sup>

## 概要

点群投影画像に基づく 3D LiDAR センサを用いた歩容認証モデルを提案し、計測距離や歩行方向の変化に対する頑健性を示す。提案する歩容認証モデルでは、3次元の歩行者点群を固定視点に投影し、歩行方向による歩容形状の変化に対する悪影響を減らす。また、複数の解像度を用いる畳み込み処理を使用し、計測距離による点群密度の変化に対応する。2D attention 機構を用いて人物の識別により有用な特徴量に着目することで、計測距離や歩行方向の異なる様々な計測環境に対して適応的に学習を行う。実験では従来手法との識別性能を比較し、提案モデルの有効性を示す。

## 1. はじめに

歩容認証は、歩く姿から個人を識別する生体認証の一種である。特に、対象者に特別な動作を行わせる必要なく、遠隔から人物を特定できる点から、大規模な犯罪捜査や広域管理システムなどに有望な個人認証手段として活発に活用されている。歩容認証分野ではカメラ映像が主な対象となっているが、一方で周囲環境を3次元の点群データとして計測できる 3D LiDAR センサが認証手段として用いられた例が少ない。LiDAR センサは、デブスカメラより高い精度や照明条件の影響に強い点など、屋外環境に対する計測に向いているものの、通常のカメラと比べて空間分解能が疎であることが原因として考えられる。我々は従来手法 [2,6] に対し、3次元の歩行者点群を固定された視点に投影することで正面と背面の2視点の歩容形状に変換し、解像度の異なる2つの2D畳み込み処理を用いることで、計測距離による点群密度の変化に対応する識別手法を提案してきた [1]。さらに本研究では、入力データのうち有用な

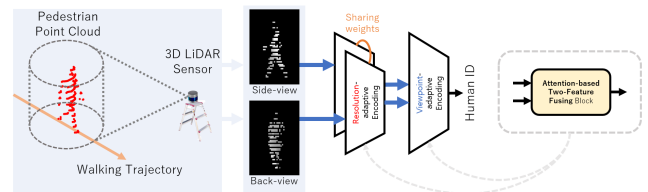


図 1 2視点の点群投影画像と 2D attention 法に基づく 3D LiDAR を用いた歩容認証の提案モデル

情報に注目する注意機構 (attention) を既存モデル [1] の改良に用いることで、LiDAR センサの照射方向・計測距離による自己オクルージョン・点群密度の変化に対してより適応的に学習できる識別モデルを開発する。

## 2. 提案手法

### 2.1 点群投影画像

本研究では、人物領域から抽出された時系列の歩行者点群のみを入力に用いる。3次元の点群データの仰俯角・方位角をそれぞれ縦軸・横軸とする距離画像に変換する従来手法と比べ、提案モデルでは3次元のユークリッド空間  $(x, y, z)$  に対して右側面と背面の2視点から深度情報を有する歩容画像列を抽出する。固定視点 (側面・背面) の推定方法や、Z-buffer 法を用いた提案画像の深度値の計算に関して、既存モデル [1] と同様である。

### 2.2 識別ネットワーク

提案ネットワークの概要を図2に示す。提案する識別ネットワークは2視点のユニットを含んだ Viewpoint-adaptive Encoding (VE) と全結合層から成り、各ユニットでは2つの空間解像度を用いた畳み込み処理から空間特徴を抽出する Resolution-adaptive Encoding (RE) と 2D ConvLSTM [5] から時間特徴を抽出する Temporal Encoding (TE) から構成される。また、2次元のチャンネル情報からより有用な特徴を捉える Hu らの 2D attention 法 [4] を改良することで、入力される2つの特徴量に対してチャンネルごとにスコアをそれぞれ付与したあと1つの特徴量に合わせる Attention-based Two-Feature Fusing (ATFF) ブロックを提案する。ここで、ATFF ブロックは VE と RE のそ

<sup>1</sup> 九州大学

<sup>2</sup> NASA / Caltech JPL

a) ahn@irvs.ait.kyushu-u.ac.jp

b) k\_nakashima@ait.kyushu-u.ac.jp

c) yoshino@irvs.ait.kyushu-u.ac.jp

d) yumi.iwashita@jpl.gov

e) kurazume@ait.kyushu-u.ac.jp

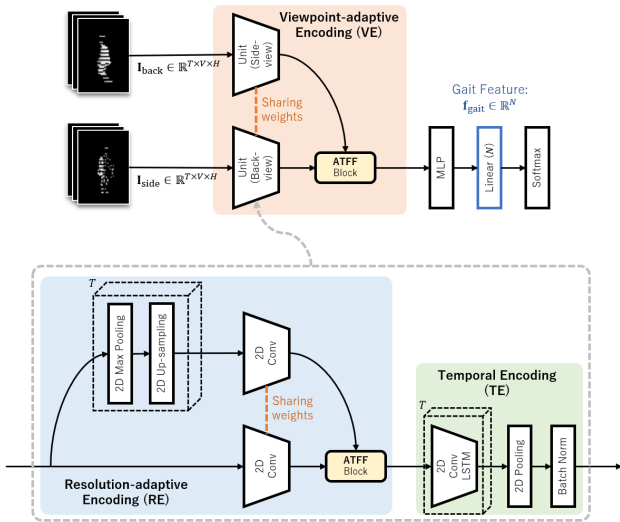


図 2 提案する識別ネットワーク (上) と単視点ユニット (下) の仕組み

それぞれのモジュールに適用されており、2つの特徴量を動的に融合する。特に、2視点の2つのユニットをそれぞれ学習させたあとそれらの特徴量を1D Pooling層で融合する既存モデル [1] と比べ、本研究での提案ネットワークでは ATFF ブロックを通して end-to-end 方式で学習を行う。

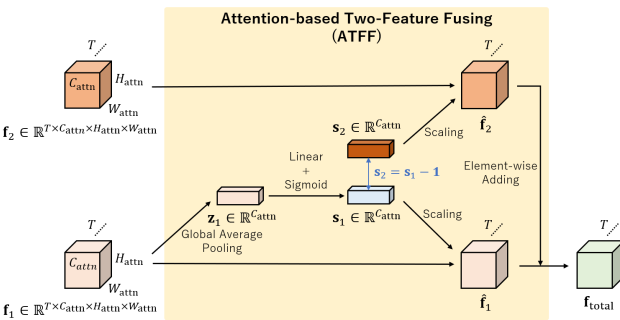


図 3 ATFF ブロックの仕組み。入力される 2 つの特徴量はチャンネルごとにスコアを付与したあと 1 つの特徴量に融合される。

### 3. 実験

#### 3.1 データセット

本実験では、我々が構築したデータセットを用いて提案手法の有効性を検証する。このデータセットでは、垂直 32 ラインの 3D LiDAR センサ (Velodyne VLP-32C) で取得された 30 人分の歩行者の点群データから構成される。具体的には、歩行者ごとに 2 つの計測距離 (10m, 20m) と 8 つの歩行方向 (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°) の 16 個の組み合わせからなる。

#### 3.2 評価方法

学習時、全歩行者 30 名のうち 20 名分の点群データを学習に使用し、残りの 10 名で評価を行った。また、本実験で

用いられる歩容画像のサイズは  $64 \times 44$  組となっており、フレーム数  $T$  は 15 と設定した。GEI [3] を用いる Benedek らの従来手法 [2] の場合、歩容画像列の深度値を二値に変換したあとフレーム  $T$  に対する平均を計算する。評価時、特徴量空間のコサイン距離に基づいて、事前に保存した辞書データ (gallery) とクエリデータ (probe) の照合を行った。

表 1 識別手法の精度評価 (%). 一位を太字で示す。

識別ネットワーク	モダリティ	点群投影方式	視点	識別率
Benedek ら [2]	GEI [3]	球面投影 [6]	変換無し	30.8
			側面	34.5
		平行投影 [1]	変換無し	38.4
			側面	58.1
Yamada ら [6]	深度画像列	球面投影 [6]	変換無し	28.7
			側面	16.3
		平行投影 [1]	変換無し	53.9
			側面	50.3
提案手法	深度画像列	球面投影 [1]	変換無し	72.2
			側面	73.4
			背面	77.8
			側面+背面	<b>81.3</b>

#### 3.3 識別性能の定量的評価

実験結果を表 1 に示す。2つの計測距離 (cross-view) と 8つの歩行方向 (cross-distance) の組み合わせから成る 16 個の実験設定に対する識別率の平均値を示しており、ここでそれぞれの識別率に対する probe は gallery の計測距離・歩行方向を含まない。比較手法 [2,6] において、線形補間なしで歩行者点群を実空間上に投影する提案投影方式が識別性能に優れている。また、Yamada らの識別手法 [6] を除いて、側面・背面などの固定視点からの歩容形状に変換することで、歩行者の動き変化の一貫性を与えて識別性能を向上させたと考えられる。最後に提案手法では、2視点の歩容形状からの歩容特徴を融合し、単視点のみの識別率を改善している。

#### 4. まとめと今後の予定

本稿では、2D attention 法を用いて空間解像度と視点の異なるそれぞれの 2 つの歩容特徴を動的に融合および学習する 3D LiDAR の歩容認証モデルを提案し、従来手法との性能比較を行った。今後は、歩行者の体型・姿勢をパラメータ化するモデルベースの識別モデルに着目し、さらに自己オクルージョンや点群密度変化に対する頑健性向上に取り組む。

#### 謝辞

本研究では、JST 次世代研究者挑戦的研究プログラム JPMJSP2136, JSPS 科研費 JP20H00230 の助成を受けたものである。

参考文献

- [1] Ahn, J., Nakashima, K., Yoshino, K., Iwashita, Y. and Kurazume, R.: 2V-Gait: Gait Recognition using 3D LiDAR Robust to Changes in Walking Direction and Measurement Distance, *Proceedings of the IEEE/SICE International Symposium on System Integration (SII)*, pp. 602–607 (2022).
- [2] Benedek, C., Gálai, B., Nagy, B. and Jankó, Z.: Lidar-Based Gait Analysis and Activity Recognition in a 4D Surveillance System, *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, Vol. 28, No. 1, pp. 101–113 (2018).
- [3] Han, J. and Bhanu, B.: Individual recognition using gait energy image, *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, Vol. 28, No. 2, pp. 316–322 (2006).
- [4] Hu, J., Shen, L. and Sun, G.: Squeeze-and-Excitation Networks, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7132–7141 (2018).
- [5] Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. and Woo, W.: Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, *Proceedings of the Neural Information Processing Systems (NIPS)*, pp. 802–810 (2015).
- [6] Yamada, H., Ahn, J., Mozos, O. M., Iwashita, Y. and Kurazume, R.: Gait-based person identification using 3D LiDAR and long short-term memory deep networks, *Advanced Robotics*, Vol. 34, No. 18, pp. 1201–1211 (2020).