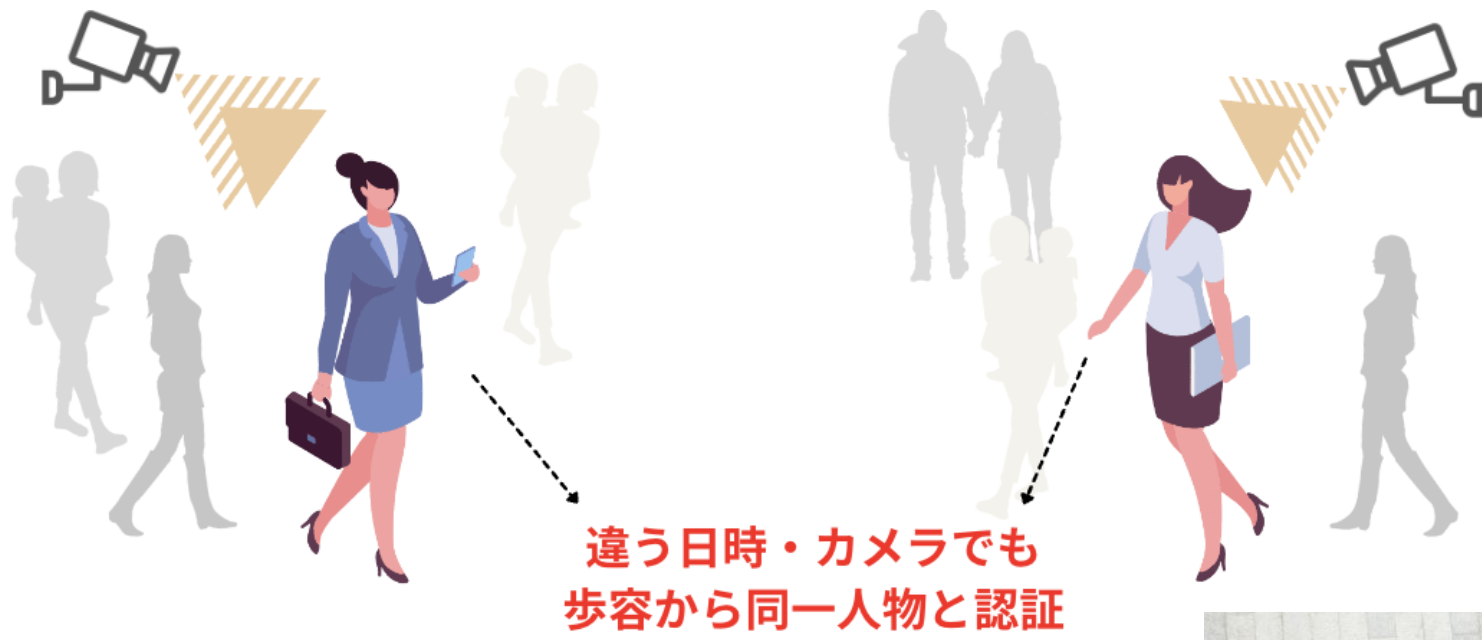


# 多視点下における3次元関節座標の歩容特徴による個人識別の精度向上

CS61c 小林研究室 256X026X 河野 瞭人 指導教員：小林 太

## 【研究背景】

### ■ 歩容による個人識別システム



### ■ 応用例

- 防犯システム
- 商業施設等でのマーケティング

## 【関連研究】

### ■ アピアランスベース[1]

- 服装や持ち物が影響
- 計算コストが低い

### ■ モデルベース

- 見た目の変化に強い
- 計算コストが高い

## ➡ 視点が変化した場合に精度が低下

## 【目的】

### ■ 座標回転により、3D 骨格データの歩行方向を統一

### ■ 骨格推定システムと単一RGBカメラのみで3D 骨格データを取得

## ➡ 多視点下において歩容撮影角度の変化にロバストな個人識別システムの開発

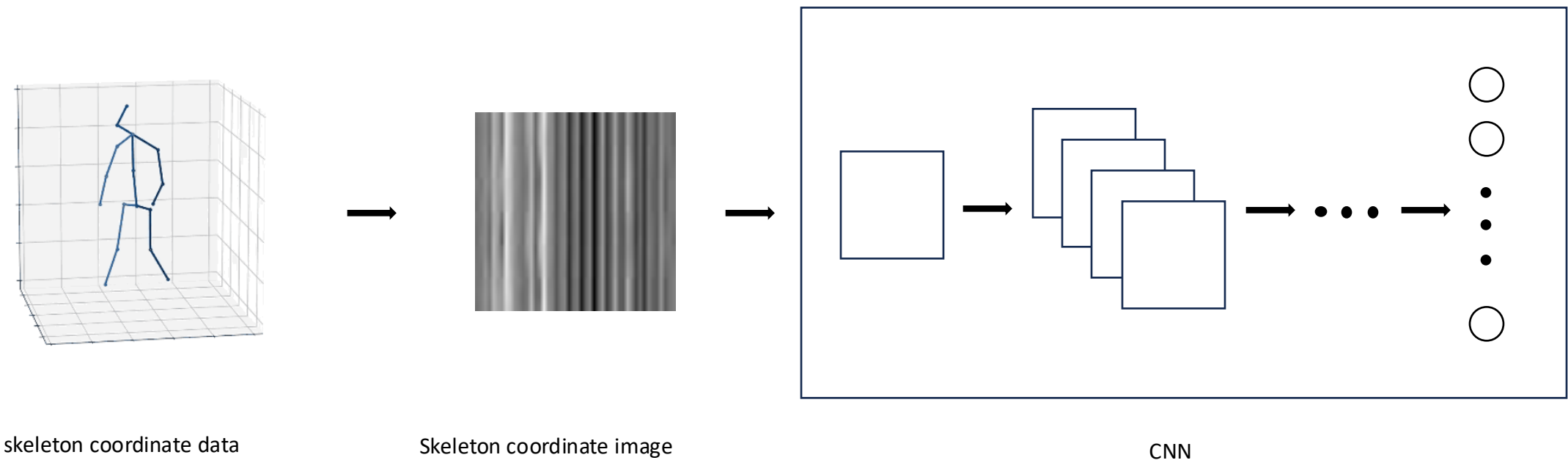
## 【歩容による個人識別】

### 個人識別システムの概要

Step1. 3D 骨格座標の取得と正規化

Step2. 骨格座標画像に変換

Step3. CNN による分類モデルで識別



### Step1. 3D 骨格座標の取得と正規化

#### ■ 3D 骨格座標の取得

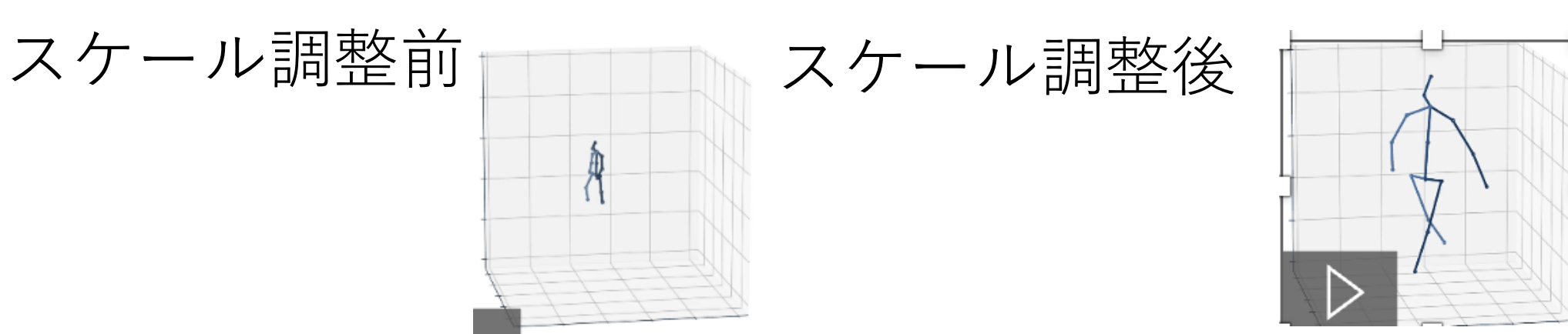
- AlphaPose[3]、MotionBERT[2]による3D骨格座標の計測

#### ■ 正規化

番号	正規化手法	処理
1	Scale-Adjustment	スケール調整
2	Viewpoint-Normalization	スケール調整 + 座標回転

#### • スケール調整

1. MidHipを原点に固定
2. MidHipからHeadの高さの差が0.95となるよう拡大・縮小



#### • 座標回転

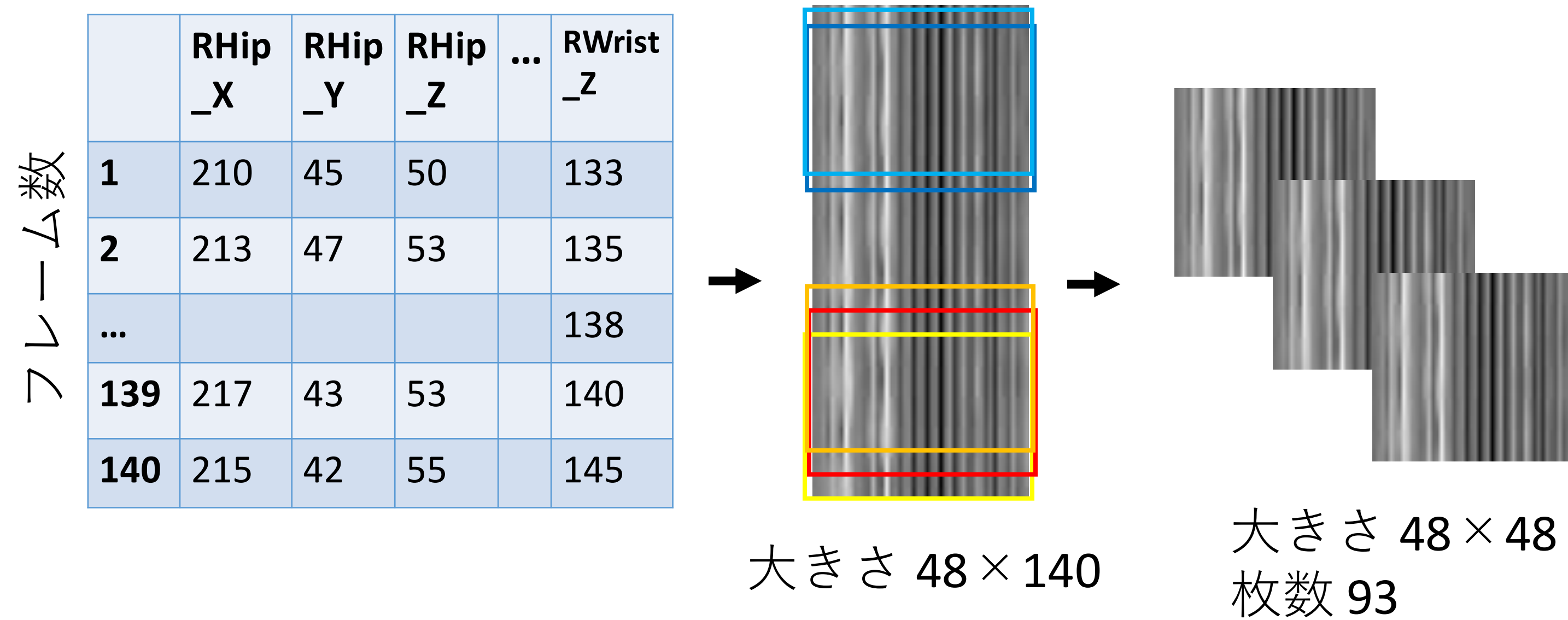
1. 2フレーム間のMidHipの差分ベクトル $v_i$ を歩行方向とする
2. 歩行方向が一定となるように座標回転



### Step2. 骨格座標画像に変換

1. 座標の値を256階調に変換
2. 歩行動画ごとに、縦軸をフレーム数、横軸を **keypoint** として1枚のグレースケール画像に変換
3. 複数枚の正方形の画像に切り取る

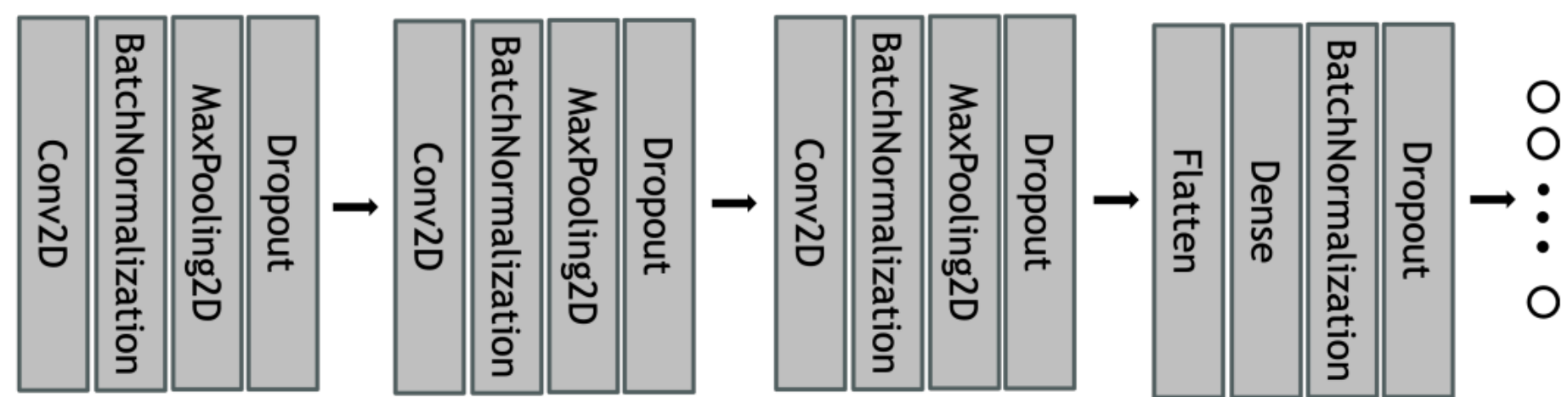
Keypoint(16個) × 3(x,y,z座標)



### Step3. CNN による分類モデルで識別

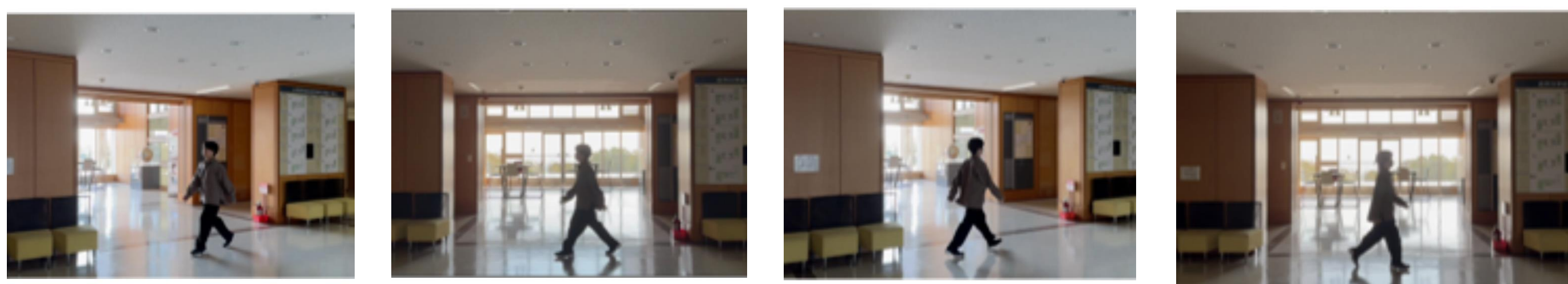
#### ■ 入力：正方形に切り取った骨格座標画像

#### ■ 出力：歩容が該当する者のラベル



## 【実験】

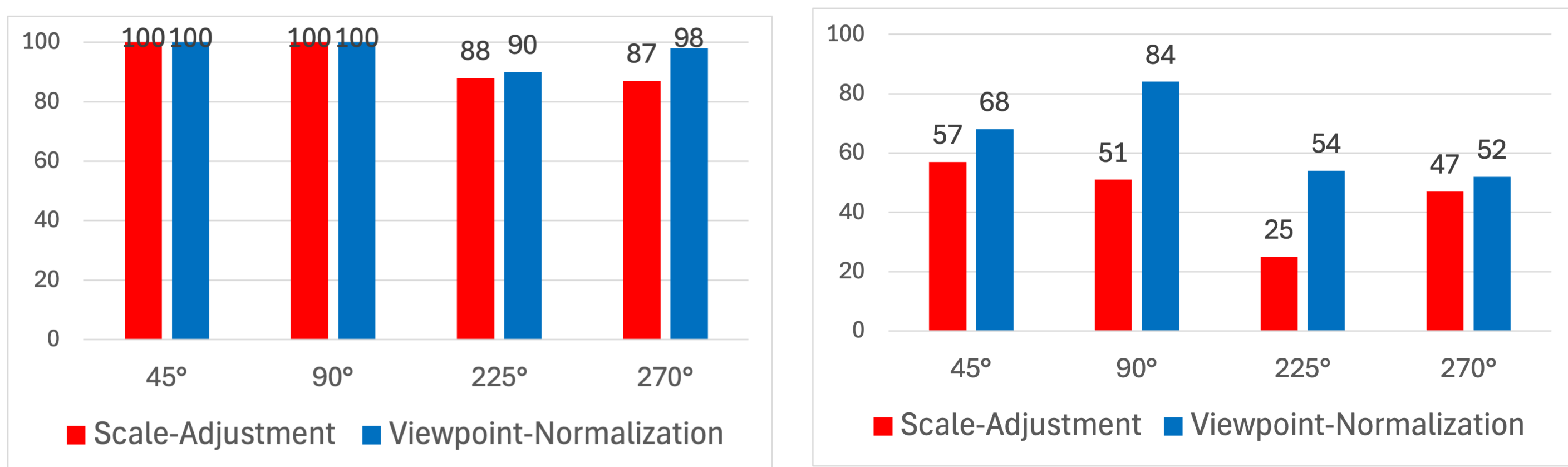
### ■ 歩行動画データ：10人 × 4方向 × 5本



### ■ データセット

1. 学習済みの視点からのデータをテストデータとする
  - 学習データ：10人 × 4方向 × 4本
  - テストデータ：10人 × 4方向 × 1本
2. 未学習の視点からのデータをテストデータとする
  - 学習データ：10人 × 3方向 × 4本
  - テストデータ：10人 × 1方向 × 5本

## 【結果と考察】



学習済みの撮影角度からのデータをテストデータとする場合の精度結果

未学習の撮影角度からのデータをテストデータとする場合の精度結果

### ■ "Viewpoint-Normalization"の手法の精度が高い

→ 座標回転の処理が有効

### ■ 45度、90度のデータの精度が225度、270度より高い

→ 45度から撮影したデータの方が体の正面を捉えるため、225度より骨格推定の精度が高い

## 【参考文献】

- [1] Han, J. and Bhanu, Bir, "Individual recognition using gait energy image", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 2, pp. 316-322 (2006)
- [2] Wentao Zhu, Xiaoxuan Ma, Zhaoyang Liu, Libin Liu, Wayne Wu, and Yizhou Wang: "MotionBERT: A Unified Perspective on Learning Human Motion Representations", IEEE Conference on Computer Vision, pp. 15039-15053 (2023)
- [3] Fang HS, Xie S, Tai YW, and et al: "RMPE: Regional multi-person pose estimation[C]", the IEEE International Conference on Computer Vision, pp. 2334-2343 (2017)