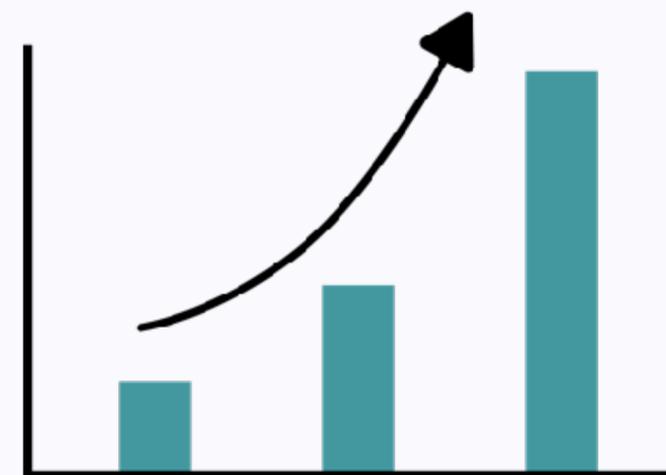
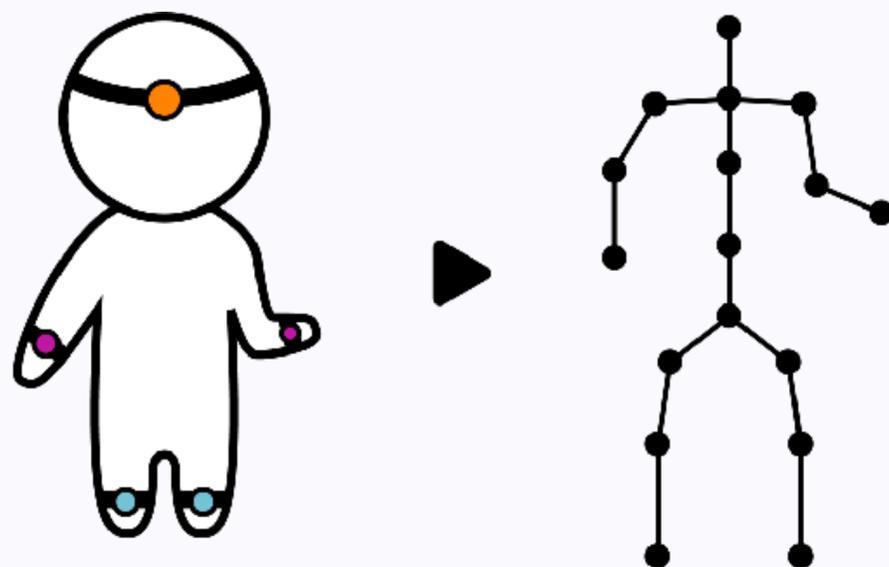


モーションキャプチャで得られる
3次元骨格情報を用いた
調理行動の分類に関する基礎検討

愛知工業大学 情報科学部 学部2年
多田隆人 梶克彦

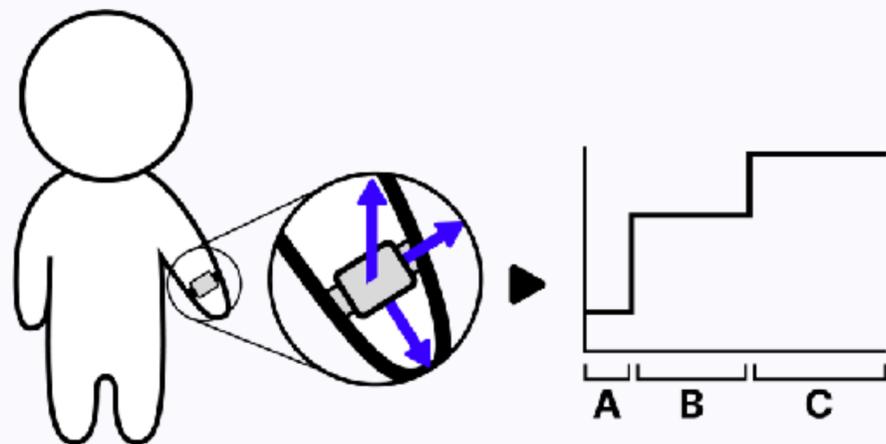
モーションキャプチャにより
手軽に骨格情報を得られる

行動認識により
技能向上の支援が期待できる



全身の骨格情報を基にした
より複雑な動作の技能向上の支援に繋がる

ウェアラブルデバイスを用いて 行動分類する研究



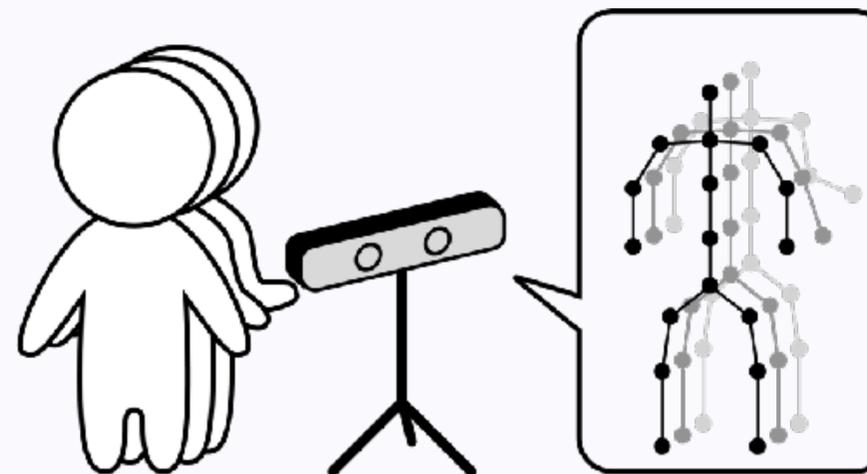
腕先の動作から行動を認識



モーションキャプチャを用いて
全身の動作から行動を認識

手首装着型の加速度センサを用いた実時間調理行動認識手法の実現, 大神・飛田
ウェアラブルセンサによるモノを用いた行動の認識について, 前川・卓也

モーションキャプチャから 人の動作を分析する研究



光学式,機械式,RGB-Dセンサなどから
3次元的な動作を取得



狭いキッチンでも扱い易い
慣性式を用いる

モーションキャプチャを用いた行動認識におけるマーカー身体対応付け作業の削減, 武田・Paula・大北・井上・出野
マーカーレスモーションキャプチャによる三次元動作解析の応用例, 春名・昆・稲垣・佐藤
モーションキャプチャを用いた演奏動作計測, 三戸

最終目標: 調理の技能評価や改善
行動認識と位置推定と手順書を組み合わせた調理行動の分類

- 3次元の骨格情報
- 作業者の位置情報
- 行動の順序, 必要な時間

- 骨格情報 から 時間ごとに行動の確率 が得られる
- 位置情報 から 場所ごとの正しい行動 が絞り込める
- 順序と時間 から 時系列における正しい行動 が絞り込める

→ 相補的に用いた絞り込みで行動認識精度の向上を目指す

目的

モーションキャプチャを用いた調理行動の分類と精度向上

アプローチ

- ① … 骨格情報から特徴量を抽出
- ② … 機械学習の手法ごとの精度の比較
- ③ … モーションキャプチャと手順書の有用性の考察

チャーハンの調理行動を対象とする

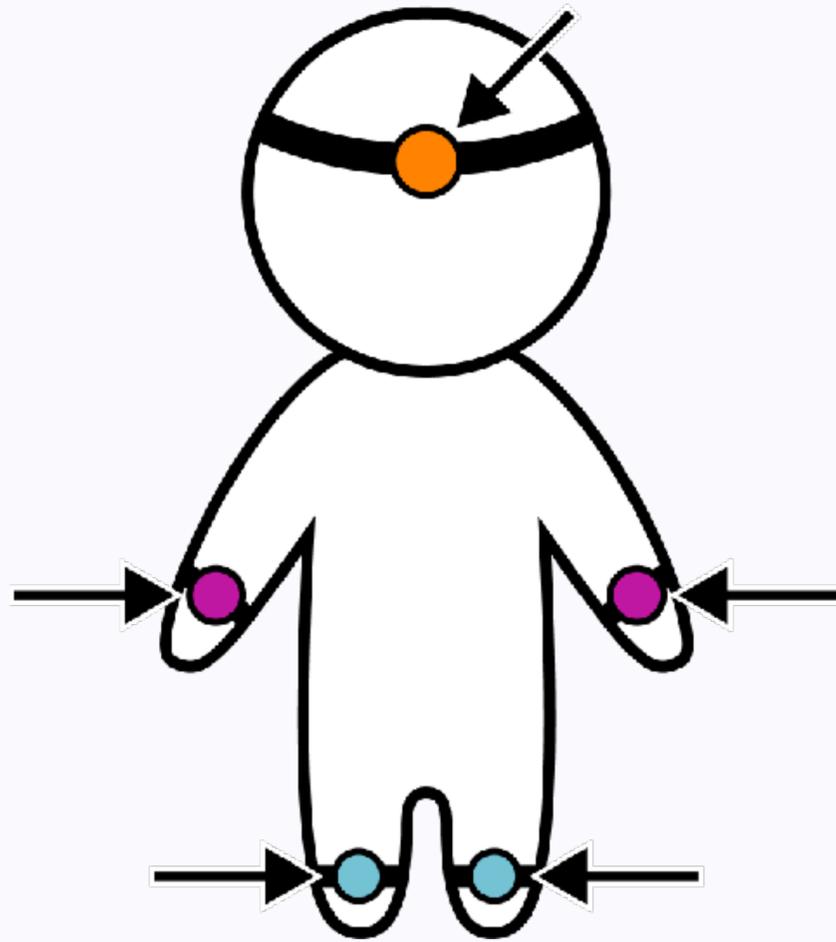
チャーハンの調理に含まれる行動

- ねぎを洗う
- ねぎを切る
- 卵を割る
- 卵を混ぜる
- 油を引く
- 卵を流し込む
- 炒める
- 米とねぎを入れる
- 味付ける
- その他

選定理由

行動分類の有効性を包括的に検証が可能

調理台, コンロ前, 流し台 など動作を伴う
洗う, 混ぜる, 炒める など多様な動作を含む



慣性式モーションキャプチャのmocopiを使用

頭・両腕・腰・両足の計6箇所にセンサを装着

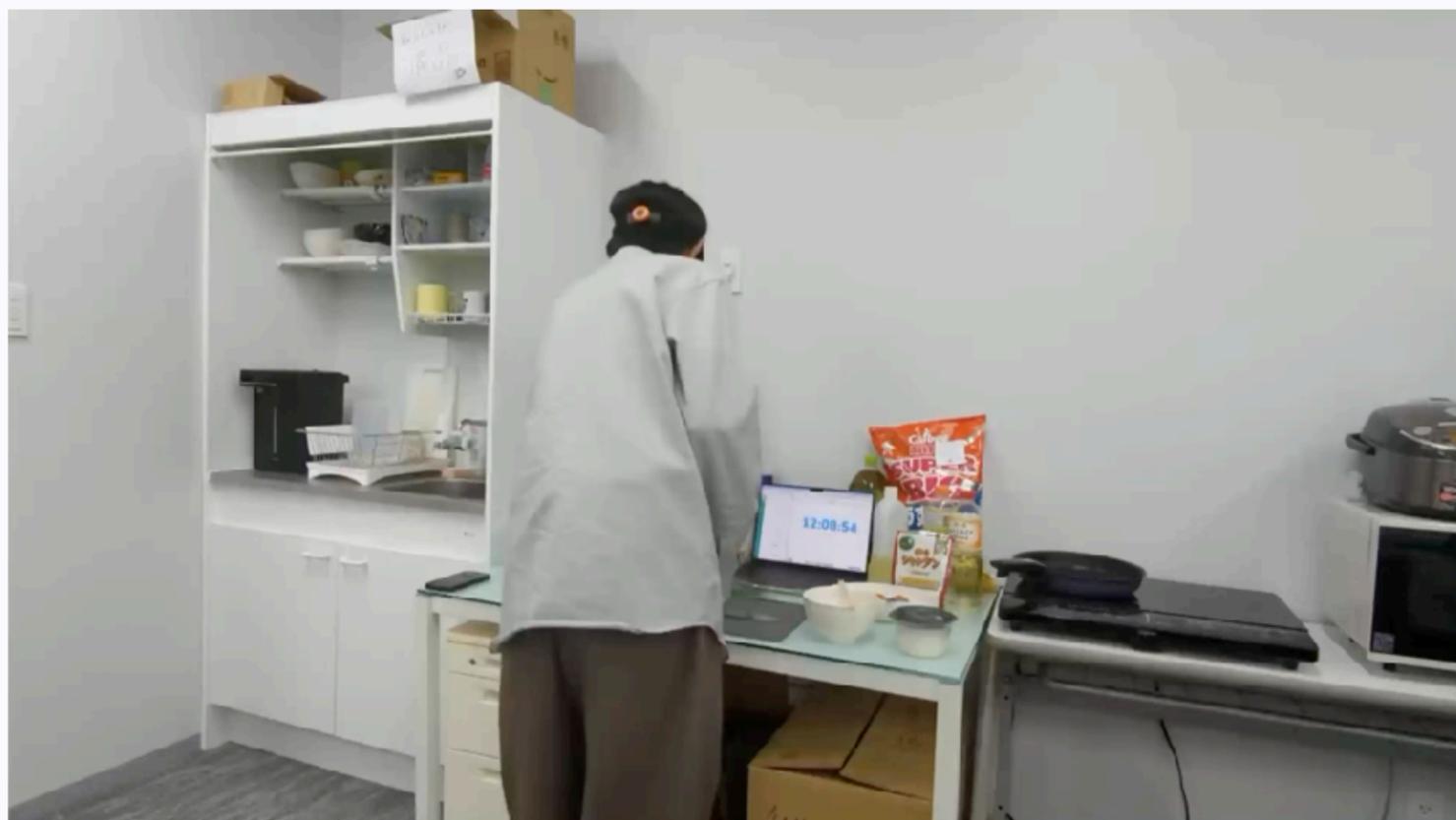
時間ごとの骨格情報と相対的な位置情報を取得可能

5人の調理行動を各2回センシング

4人分を学習データ, 1人分をテストデータとする

動画をもとに手動で正解データを作成

実際の映像

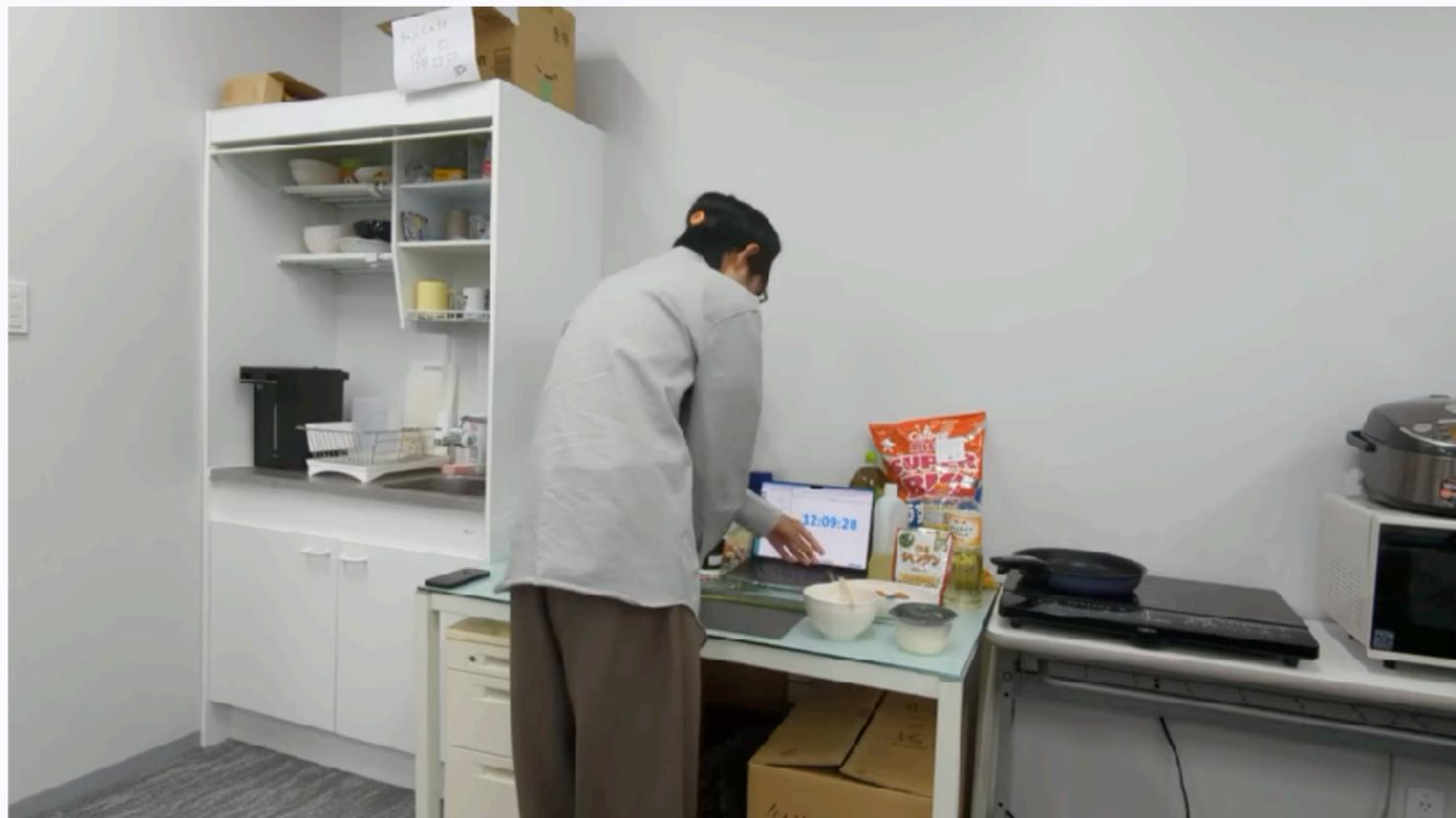


取得した骨格情報



ねぎを洗う / 移動

実際の映像

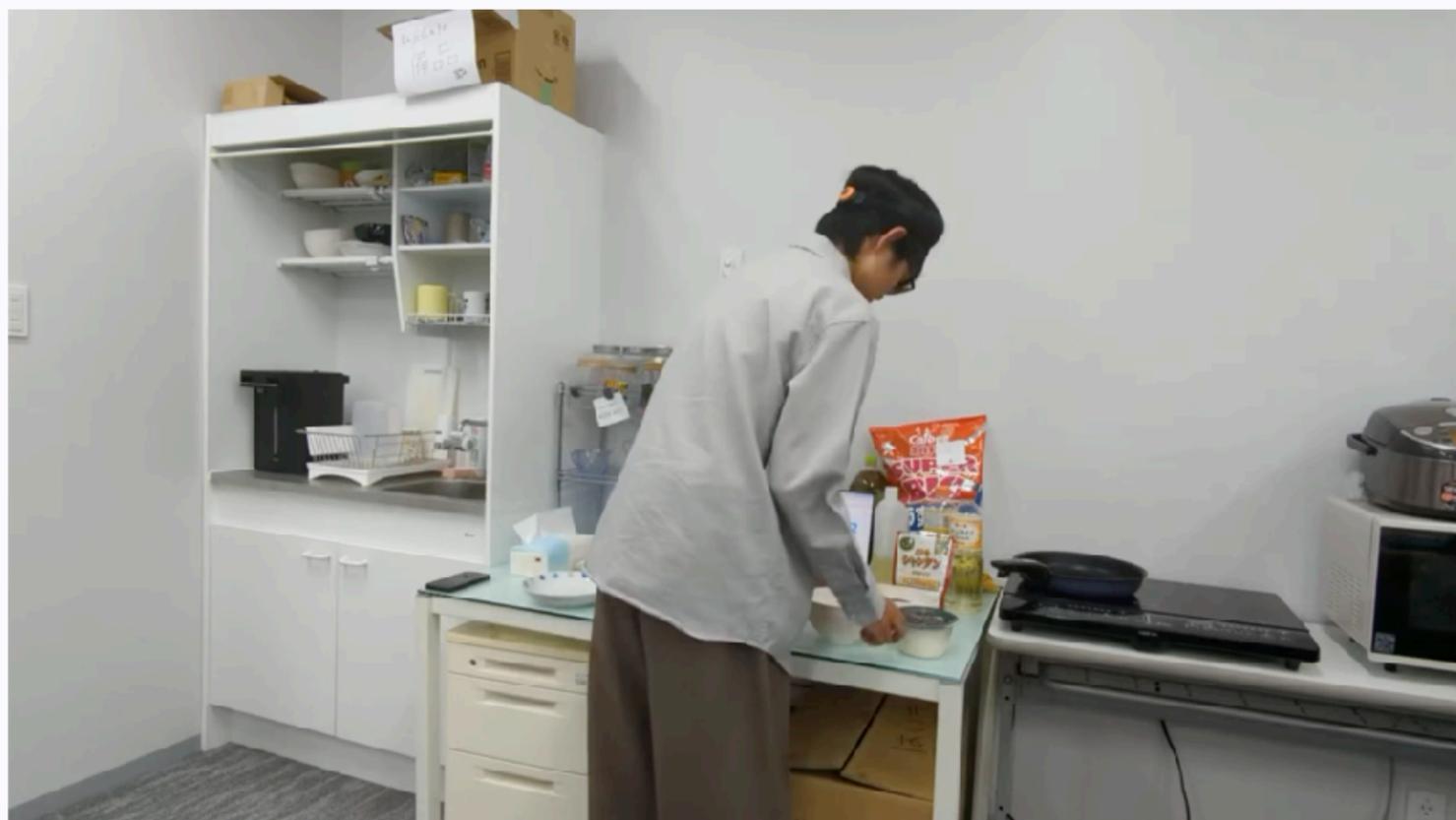


取得した骨格情報



ねぎを切る

実際の映像

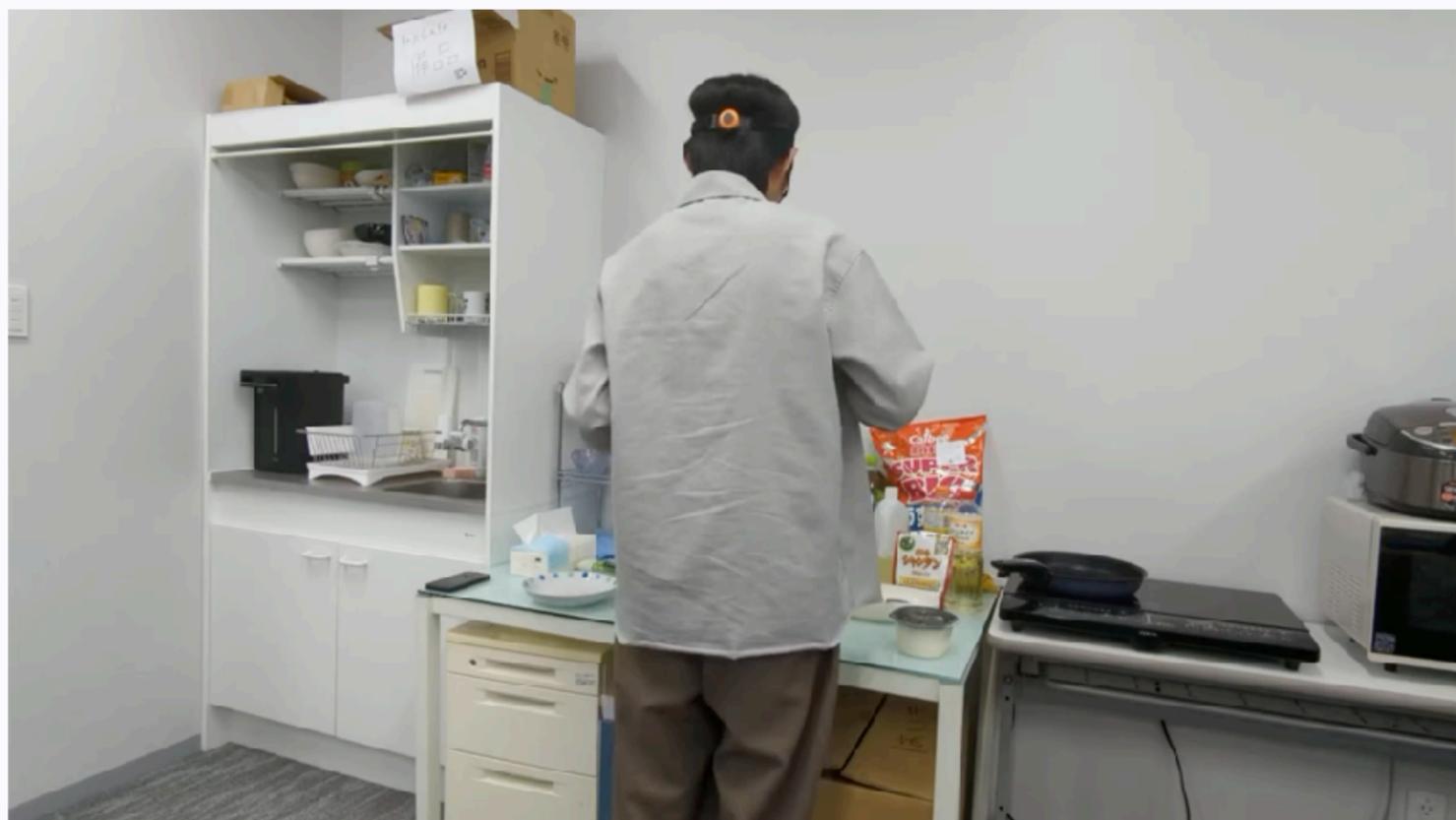


取得した骨格情報



卵を割る

実際の映像

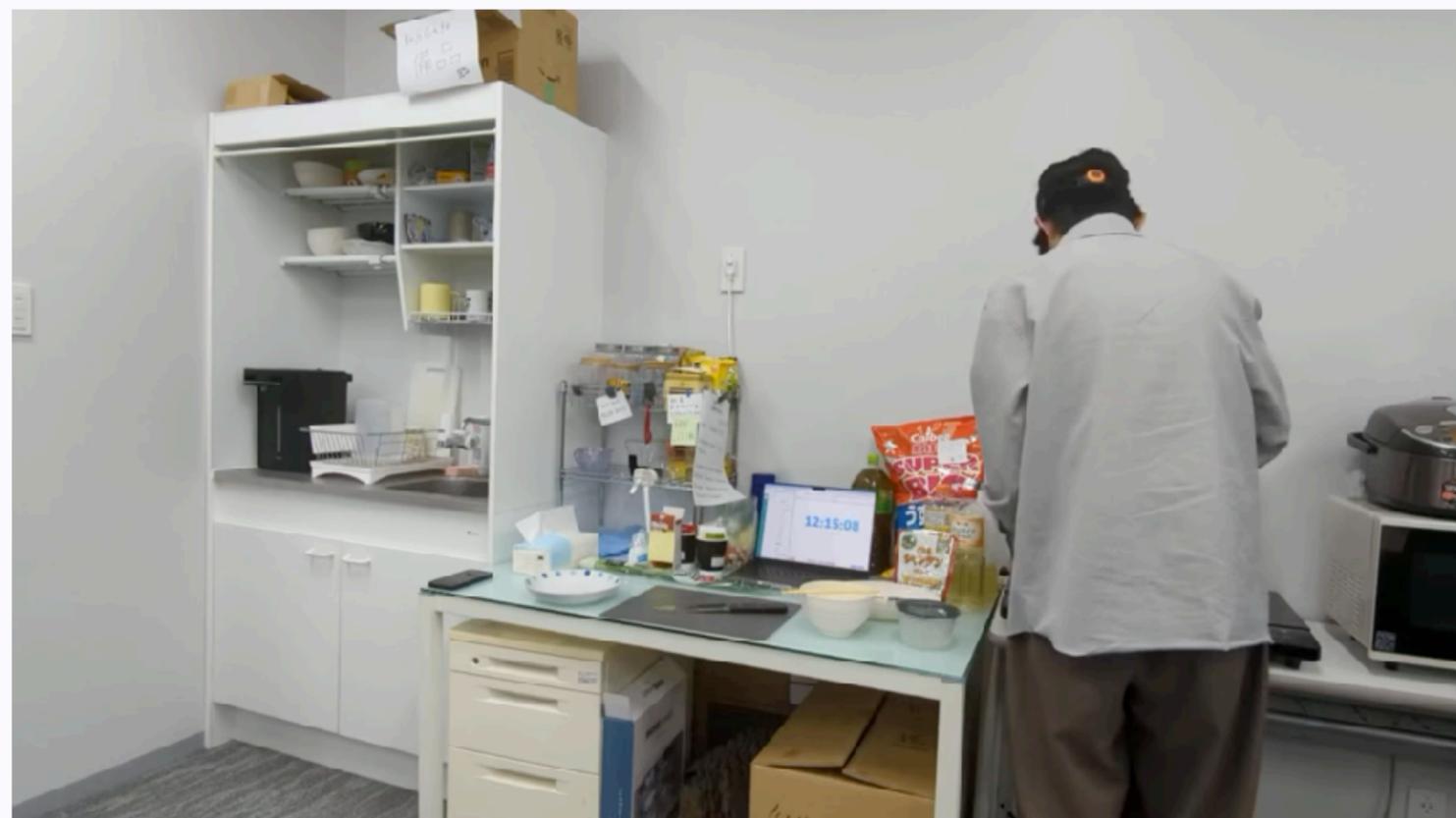


取得した骨格情報



卵を混ぜる

実際の映像

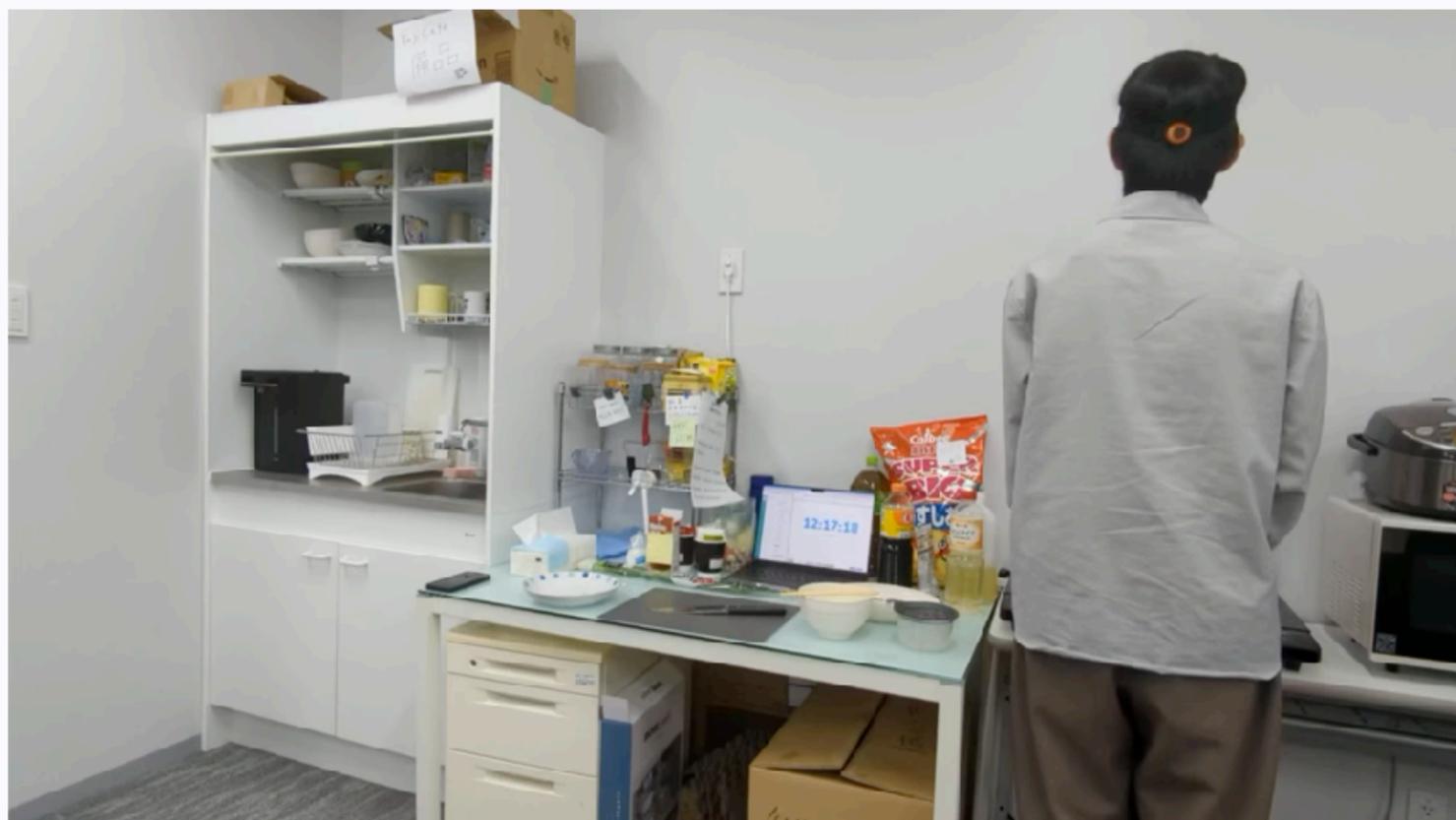


取得した骨格情報



炒める

実際の映像



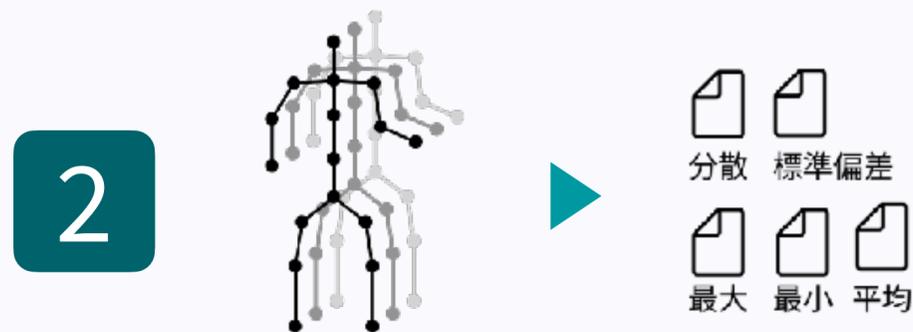
取得した骨格情報



盛り付ける



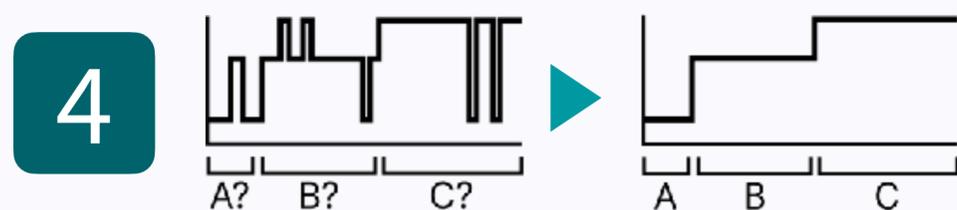
慣性式モーションキャプチャのmocopiを用いて 骨格情報(BVH形式)を得る



骨格情報から一定のウィンドウ幅 (セグメント時のウィンドウ幅)で分割し特徴量を抽出する
特徴量には最大・最小・平均・分散・標準偏差を用いる

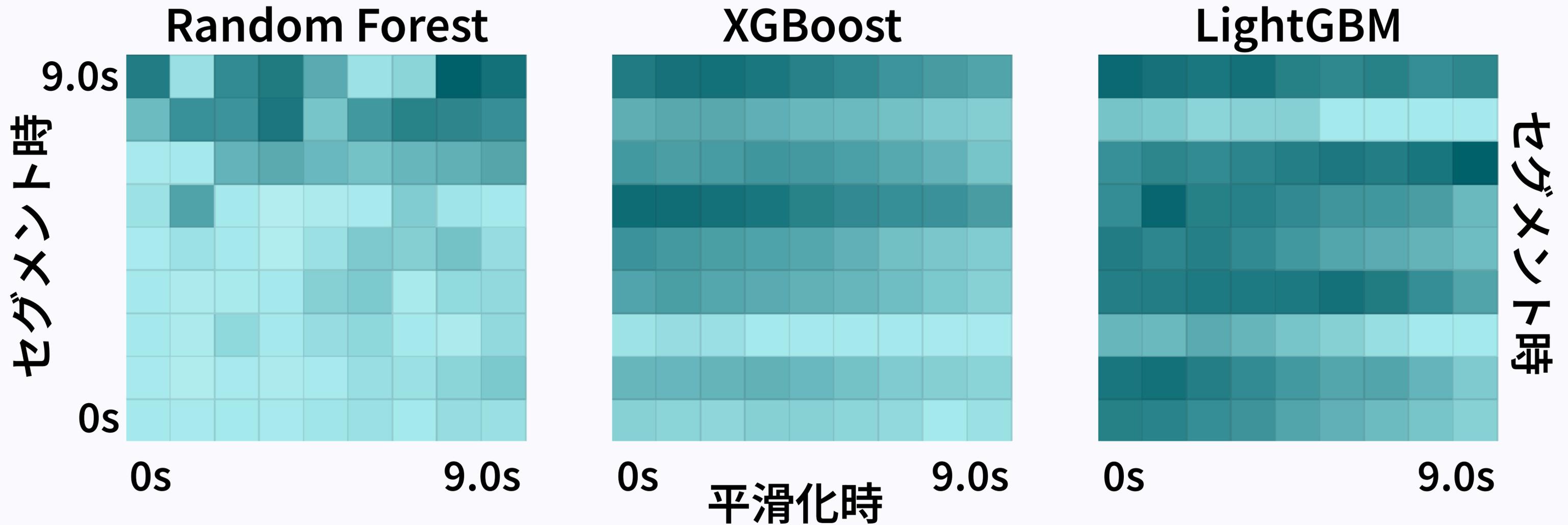


Random Forest, XGBoost, LightGBMの3つの手法を用いて分類する



分類結果を一定のウィンドウ幅 (平滑化時のウィンドウ幅)で分割し平滑化する

精度が高いほど濃い色で示す



セグメント時のウィンドウ幅が 2.0s
平滑化時のウィンドウ幅が 6.0s
のとき各手法において比較的高い精度が得られた

骨格情報からは時間ごとの行動を絞り込むことを目的とするため
Top-1に加えてTop-3でも比較

手法	Top-1 の精度	Top-3 の精度
Random Forest	0.63	0.86
XGBoost	0.81	0.96
Light GBM	0.78	0.93

XGBoost が最も精度が高く、LightGBM も比較的精度が高い
今回は **XGBoost が適している**と考えられる

XGBoost と Light GBM で
高い精度を得られた



骨格情報は特徴相互作用が複雑
XGBoost と LightGBMが
特徴相互作用を捉えるのに適切

XGBoost は LightGBM より
精度が高い

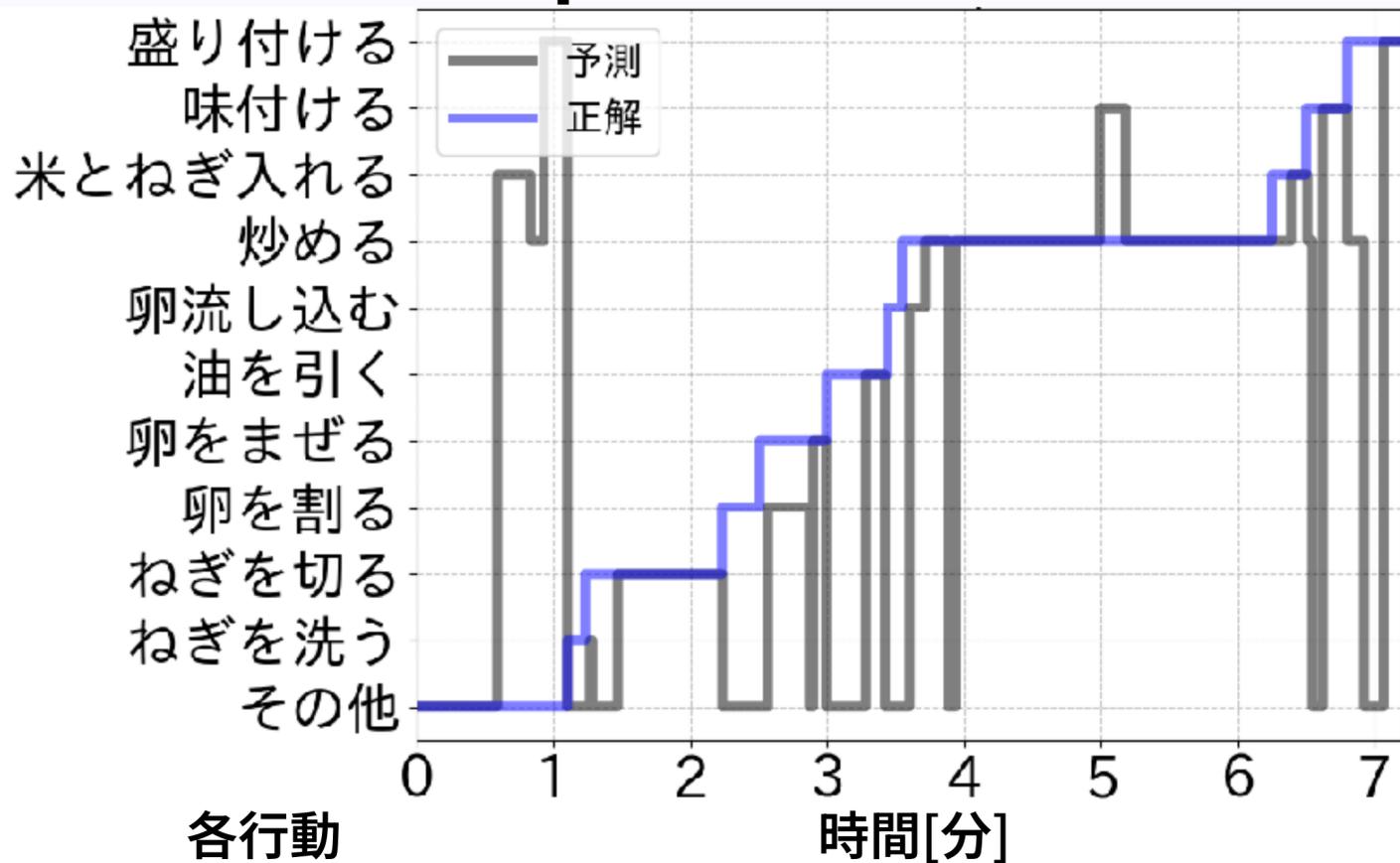


今回のデータ量では足りなく
過学習が起こったのでは

③ 評価実験 - 位置情報と手順書の精度向上への有用性の検討 18

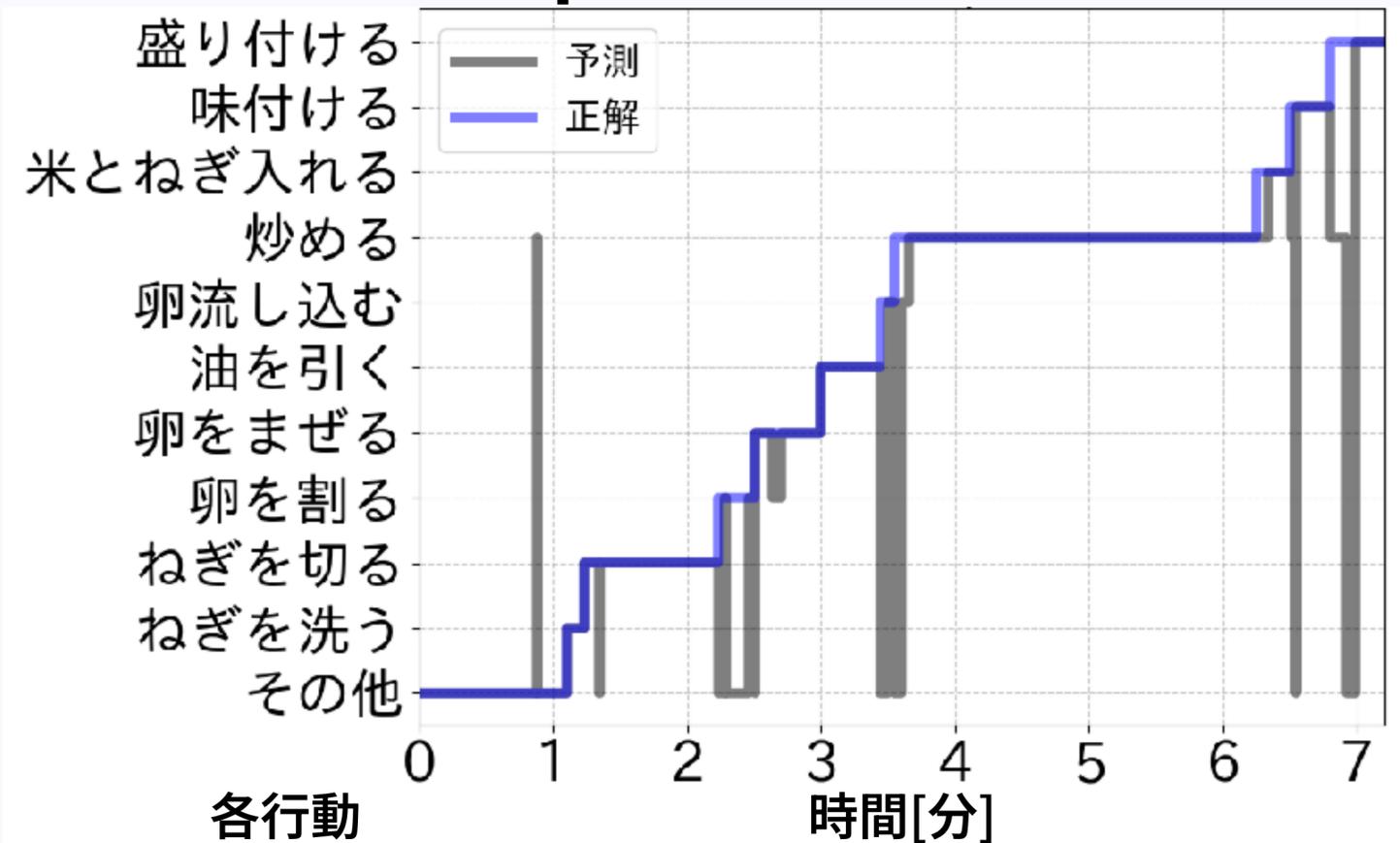
Top-3は推定結果の上位3つのうち、
正解を含む場合は正解の行動を、含まない場合は最も確率が高い行動を示した

Top-1 の結果



「ねぎを切る」「炒める」は比較的精度が高い

Top-3 の結果



Top-1で誤認識されやすい動作も
Top-3においては高い精度が得られた

周期的な動作が少ない行動は比較的精度が低い
「油を引く」「盛り付ける」



1つの行動の中で様々な特徴を含む
似た特徴を含む行動間で誤認識されやすい
このような行動も Top-3 では高い精度が得られた



位置情報や手順書を元にTop-Nから絞り込むと
精度向上が期待できる

調理行動の混同行列

	...	H	I	J	K
...	...	∴	∴	∴	∴
H.炒める979	.000	.000	.000
I.米とねぎ入れる151	.849	.000	.000
J.味付ける060	.013	.923	.000
K.盛り付ける147	.000	.000	.773

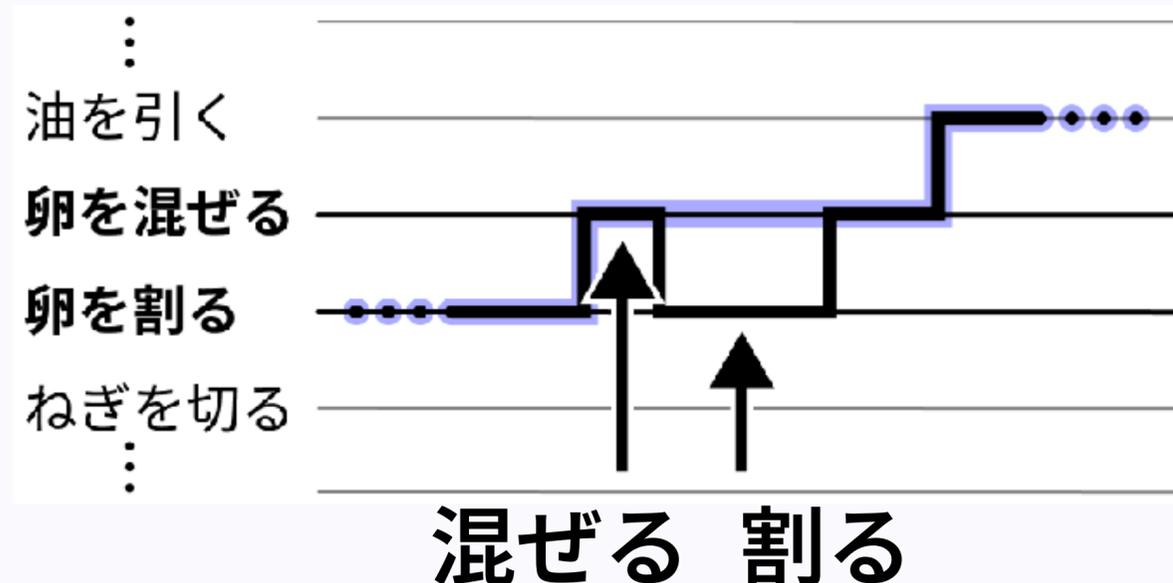
— 位置情報と組み合わせる —

「盛り付ける」が「炒める」と混同されやすい



調理台であれば「盛り付ける」
コンロ前であれば「炒める」とであると絞り込める

工程の推定結果



— 手順書と組み合わせる —

「卵を混ぜる」の後に「卵を割る」がきている



卵を混ぜる前に卵を割るという工程が必要なため
誤りであるとわかる

まとめ

モーションキャプチャで得られる3次元の骨格情報を基に調理行動を分類

今回の実験においては XGBoost が適していた

骨格情報のみでは似た動作を含む行動を誤認識しやすい

位置情報や手順書との組み合わせによる精度を向上

今後の課題

骨格情報・手順書・位置情報から相補的に行動を分類

モーションキャプチャでは人の位置情報を取得できる

調理行動においては調理工程の手順書(レシピ)を用いられる

質疑応答用

	A	B	C	D	E	F	G	H	I	J	K
A.その他	.999	.000	.000	.000	.000	.000	.000	.001	.000	.000	.000
B.ねぎを洗う	.010	.990	.000	.000	.000	.000	.000	.000	.000	.000	.000
C.ねぎを切る	.005	.000	.995	.000	.000	.000	.000	.000	.000	.000	.000
D.卵を割る	.388	.000	.005	.607	.000	.000	.000	.000	.000	.000	.000
E.卵をまぜる	.003	.000	.000	.064	.934	.000	.000	.000	.000	.000	.000
F.油を引く	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000
G.卵流し込む	.226	.000	.000	.000	.000	.000	.774	.000	.000	.000	.000
H.炒める	.010	.000	.000	.000	.000	.000	.010	.979	.000	.000	.000
I.米とねぎ入れる	.000	.000	.000	.000	.000	.000	.000	.151	.849	.000	.000
J.味付ける	.004	.000	.000	.000	.000	.000	.000	.060	.013	.923	.000
K.盛り付ける	.080	.000	.000	.000	.000	.000	.000	.147	.000	.000	.773

実際の映像



取得した骨格情報



深層学習に対するメリット

- 少量のデータでも高い精度を出せる
- 計算コストが低い
- 過学習しにくい※

行動認識と位置推定と手順書の組み合わせによる
精度向上を目標としている



今回は機械学習単体では精度よりも手軽さを求めるため
従来の機械学習アルゴリズムを採用

材料や器具の初期位置



レシピを手順書として用いる



準備は工程に含まれないため
あらかじめ用意

冷蔵庫から取り出す動作等も
認識が期待できる

冷蔵庫内の位置は固定ではない

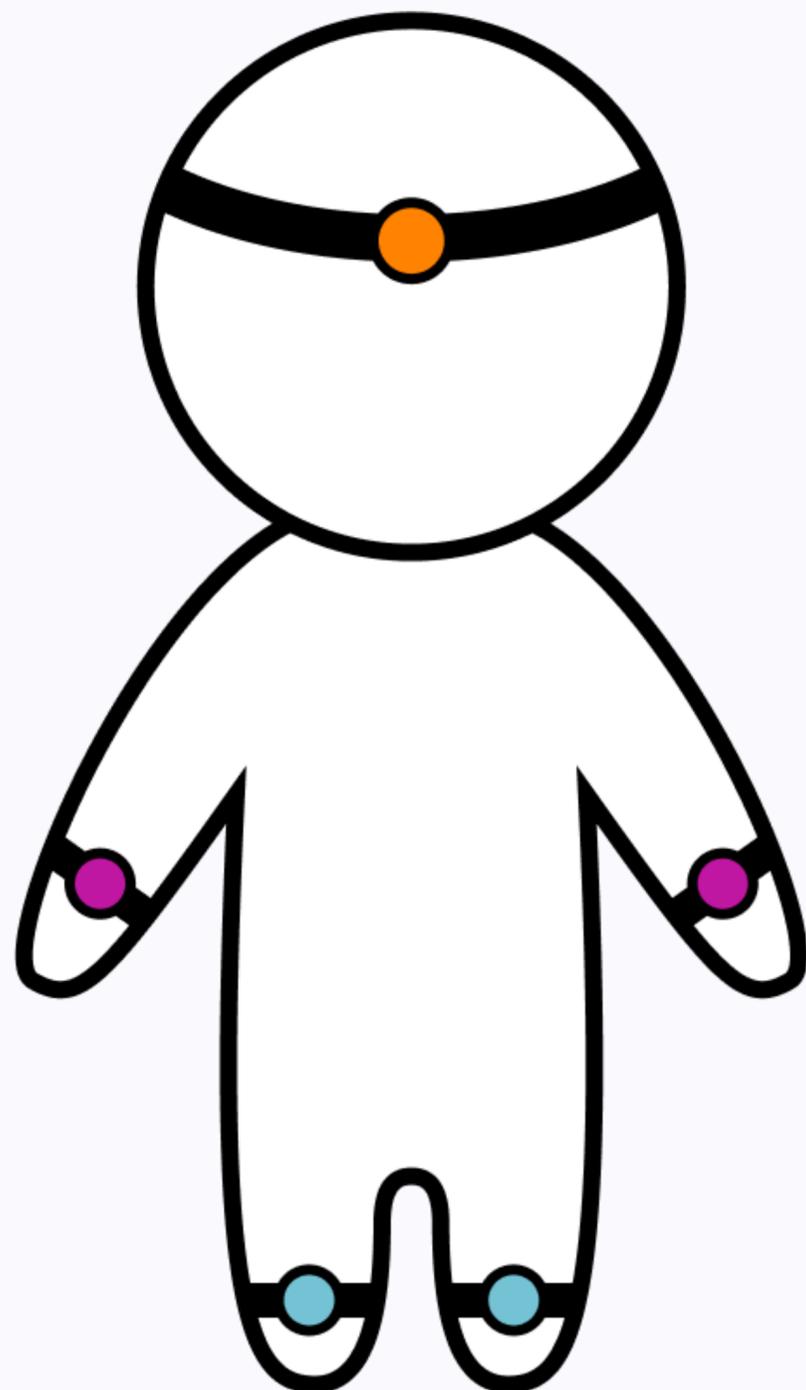


冷蔵庫の位置や開閉動作をもとに
認識できるのでは

レシピのような手順書がある行動であれば
他の行動においても期待できる

例) 工場における生産過程 (セル生産方式)

部品を取る ▶ ねじを複数箇所締める ▶ 次に渡す



頭・両腕・腰・両足の計6箇所に装着
加速度と角速度をセンシング



mocopi公式アプリ※にBluetoothで送信
アプリ側で機械学習を用いて骨格を推定



動画



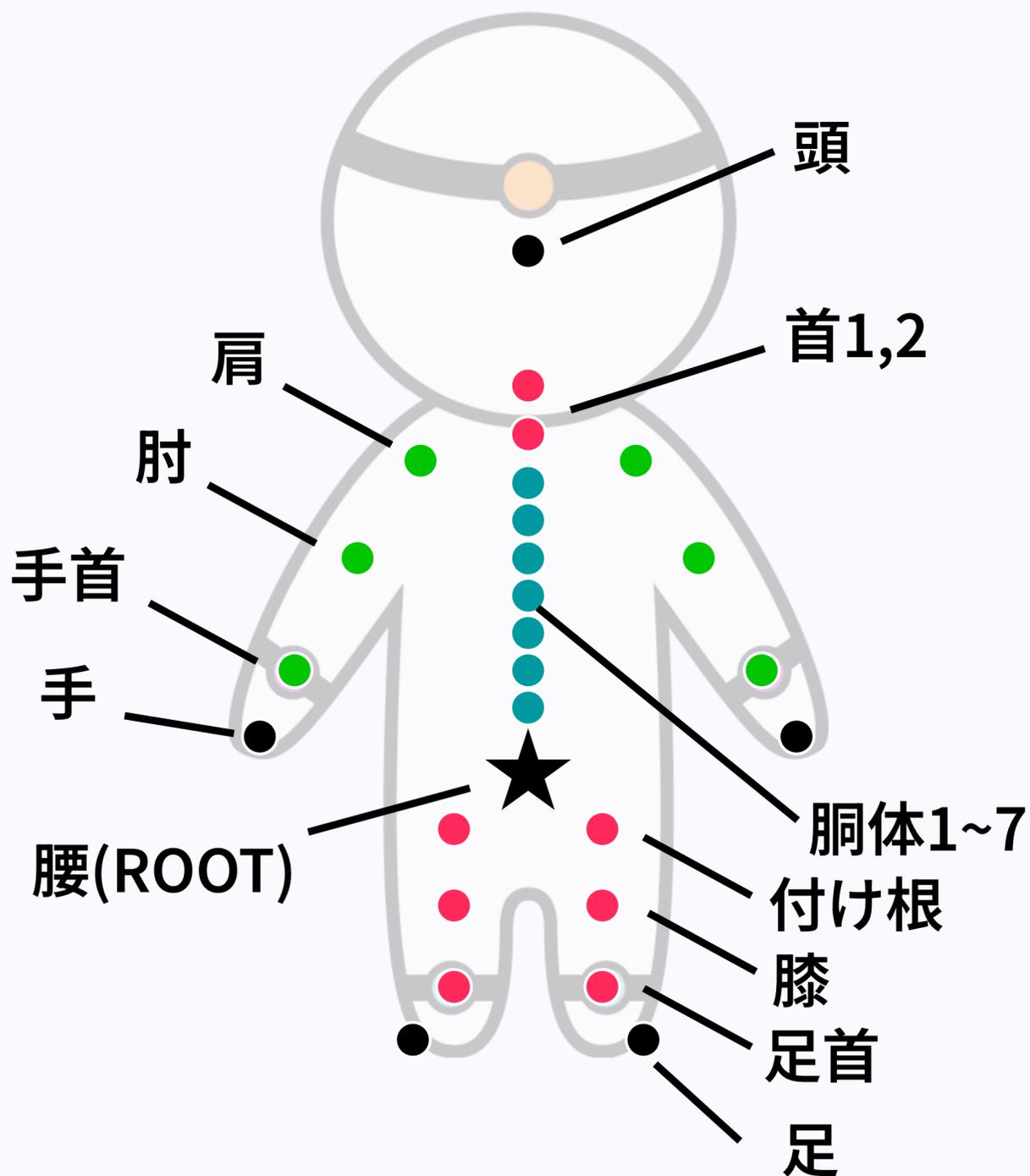
BVH形式



生データ

※骨格まで出してくれるアプリと生データを取得するだけの二種類ある

BVHファイルに含まれる関節



腰を基準として親関節からの相対的な角度と位置を示す
腰の位置は開始地点を基準とした相対的な位置を示す

※腰以外の位置は常に固定値

6箇所センサーをもとに
mocopiアプリが各関節を推定している

HIERARCHY部では骨格構造を定義している

```
1 HIERARCHY
2 ROOT root
3 {
4   OFFSET 0 93.2929 0
5   CHANNELS 6 Xposition Yposition Zposition Zrotation Xrotation Yrotation
6   JOINT torso_1
7   {
8     OFFSET 0 5.07867 -1.15138
9     CHANNELS 6 Xposition Yposition Zposition Zrotation Xrotation Yrotation
10    JOINT torso_2
11    {
12      ...
13      {
14        OFFSET -2.53307e-17 9.49737 -0.914492
15        CHANNELS 6 Xposition Yposition Zposition Zrotation Xrotation Yrotation
16        JOINT torso_7
17        {
18          ...
19          {
20            OFFSET -2.00958e-17 4.83041 0.765062
21            CHANNELS 6 Xposition Yposition Zposition Zrotation Xrotation Yrotation
22            End Site
23            {
24              OFFSET 0 0.1 0
25            }
```

MOTION部ではフレームごとの
各関節の角度と位置を示している

```
1 MOTION
2 Frames: 33520
3 Frame Time: 0.016667
4 0 93.2929 0 -0 0 -0 0 5.07867 -1.15138 -0 0 -0 0 5.61661 1.07143 -0 0 -0 -1.02604e-17 ...
5 -47.0603 92.8165 -11.3194 -3.42661 -2.56392 -108.244 0 5.07867 -1.15138 ...
6 -47.0569 92.8159 -11.3133 -3.429 -2.55492 -108.242 0 5.07867 -1.15138 ...
7 -47.0535 92.8156 -11.3069 -3.43599 -2.54376 -108.239 0 5.07867 -1.15138 ...
8 -47.0497 92.8157 -11.3006 -3.44299 -2.53261 -108.235 0 5.07867 -1.15138 ...
9 -47.0454 92.8155 -11.2947 -3.44312 -2.52127 -108.233 0 5.07867 -1.15138 ...
10 -47.0411 92.8153 -11.2887 -3.44324 -2.50994 -108.23 0 5.07867 -1.15138 ...
11 -47.0366 92.8143 -11.2821 -3.44324 -2.50993 -108.23 0 5.07867 -1.15138 ...
12 -47.032 92.8145 -11.2748 -3.46313 -2.50869 -108.237 0 5.07867 -1.15138 ...
```

処理には全てPythonを使用

BVHファイルを mcp-persor※を用いてパース



- HIERARCHY部 をオブジェクトとして取得
- MOTION部を DataFrame として取得
- 関節名一覧を取得
- 初期位置の操作
- 各情報の上書き
- CSV, BVH形式での出力 等

※mocopiのBVHファイルをパースするための自作ライブラリ (公開済)