

at-8.姿勢推定

(ディープラーニングのシステムとプログラミング)
(全12回)

<https://www.kkaneko.jp/ai/at/index.html>

金子邦彦



アウトライン

1. 姿勢推定の仕組み

姿勢推定の基本原理 人体の特徴点（キーポイント）を検出し、それらを結びつけて全身の姿勢を推定

2. 2次元と3次元の姿勢推定

2次元推定では平面上の姿勢を、3次元推定では奥行きも含めた空間的な姿勢を推定可能

3. ディープラーニングによる姿勢推定の進歩と課題

多人数や部分遮蔽への対応など精度と適用範囲が向上 訓練データ不足などの課題に対し、自己教師あり学習などの新アプローチも

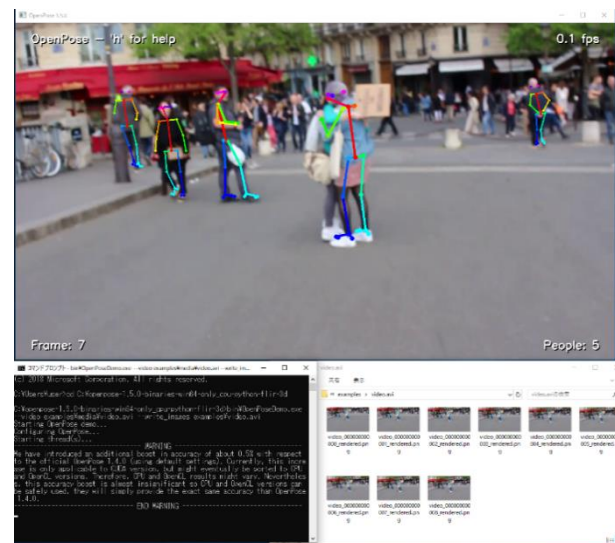
4. 姿勢推定の具体的手順

部位の位置推定 → 同一人物のキーポイント特定 → (3次元推定時) 奥行き推定

8-1. 姿勢推定の基本

姿勢推定

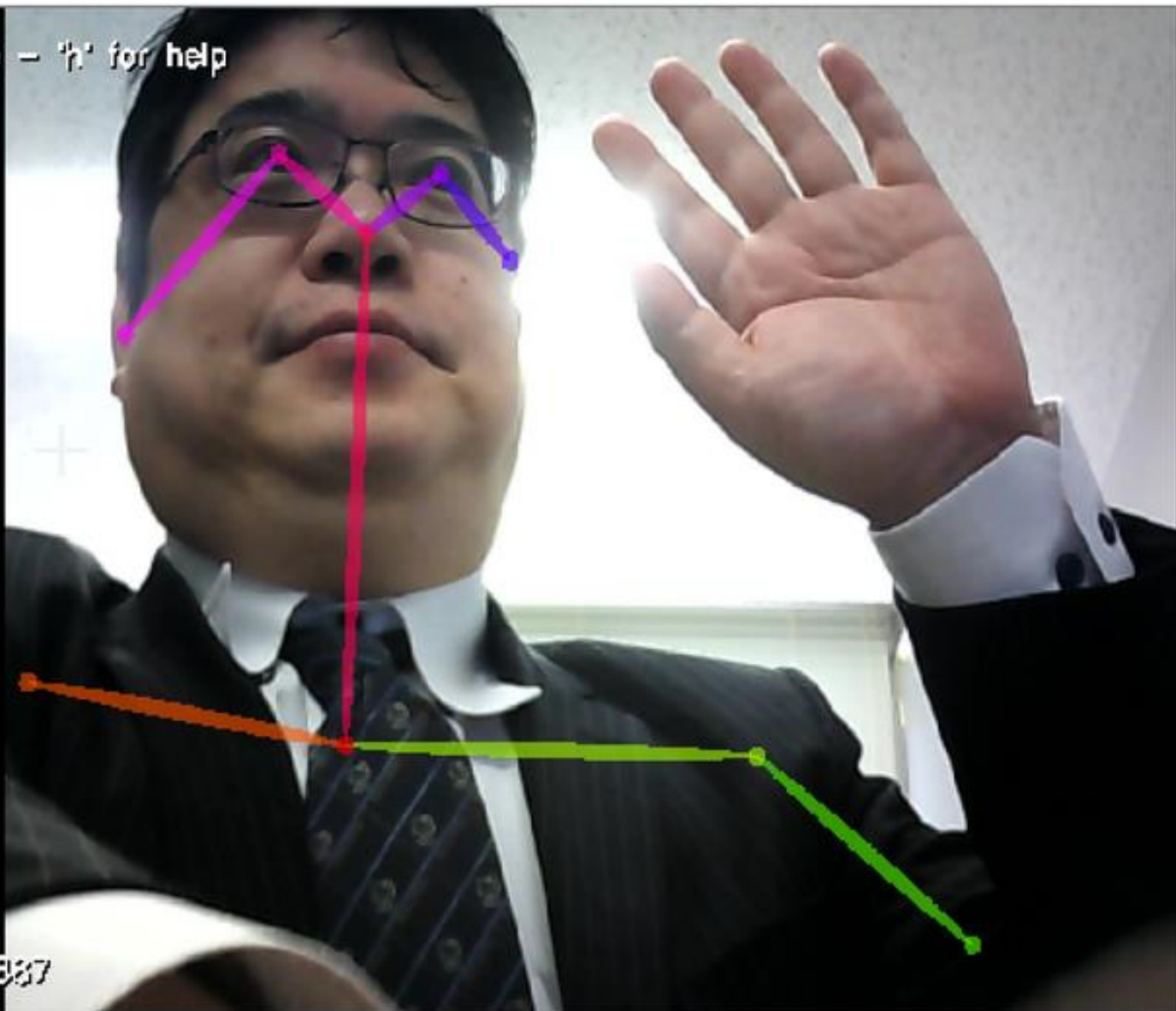
姿勢推定は、人間の全身（人体）、頭部、およびその他のオブジェクトの**位置と方向**を推定する技術



姿勢推定の結果

OpenPose - 'h' for help

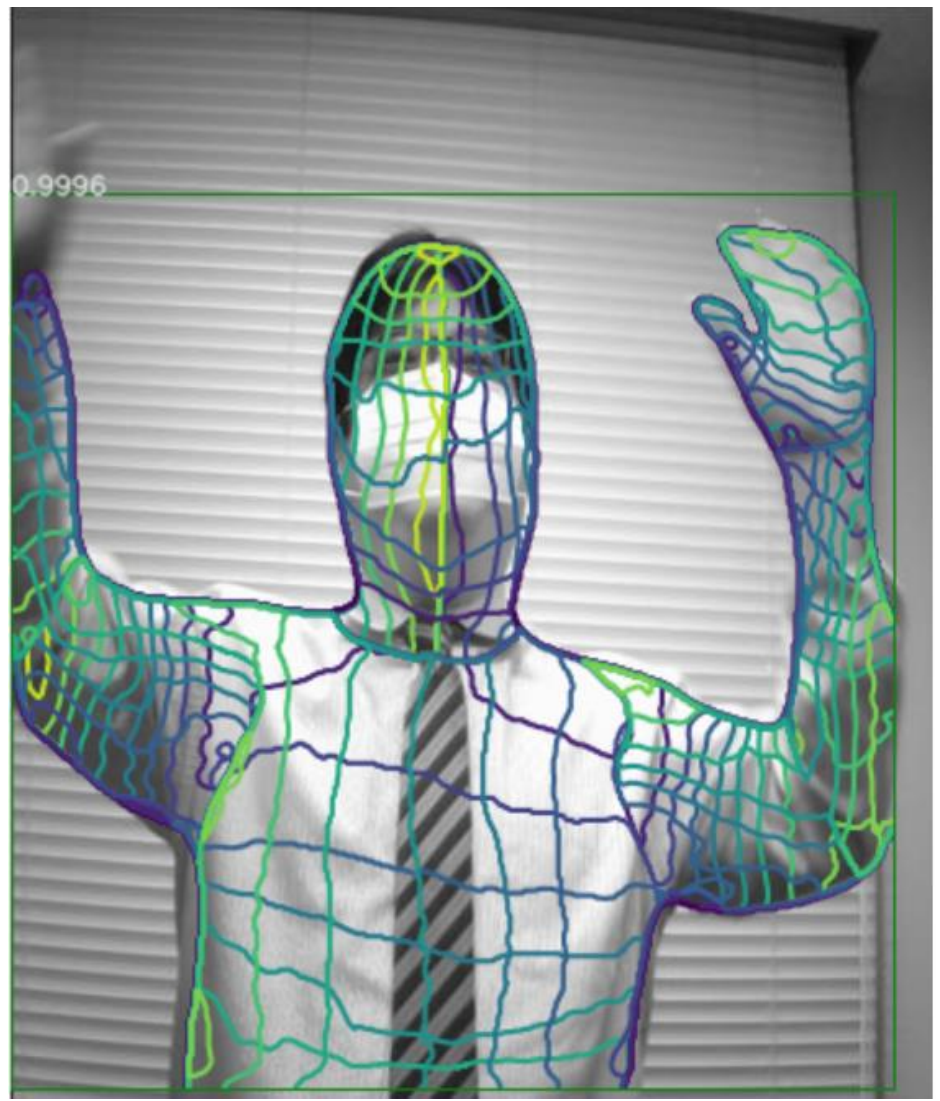
30.6 fps



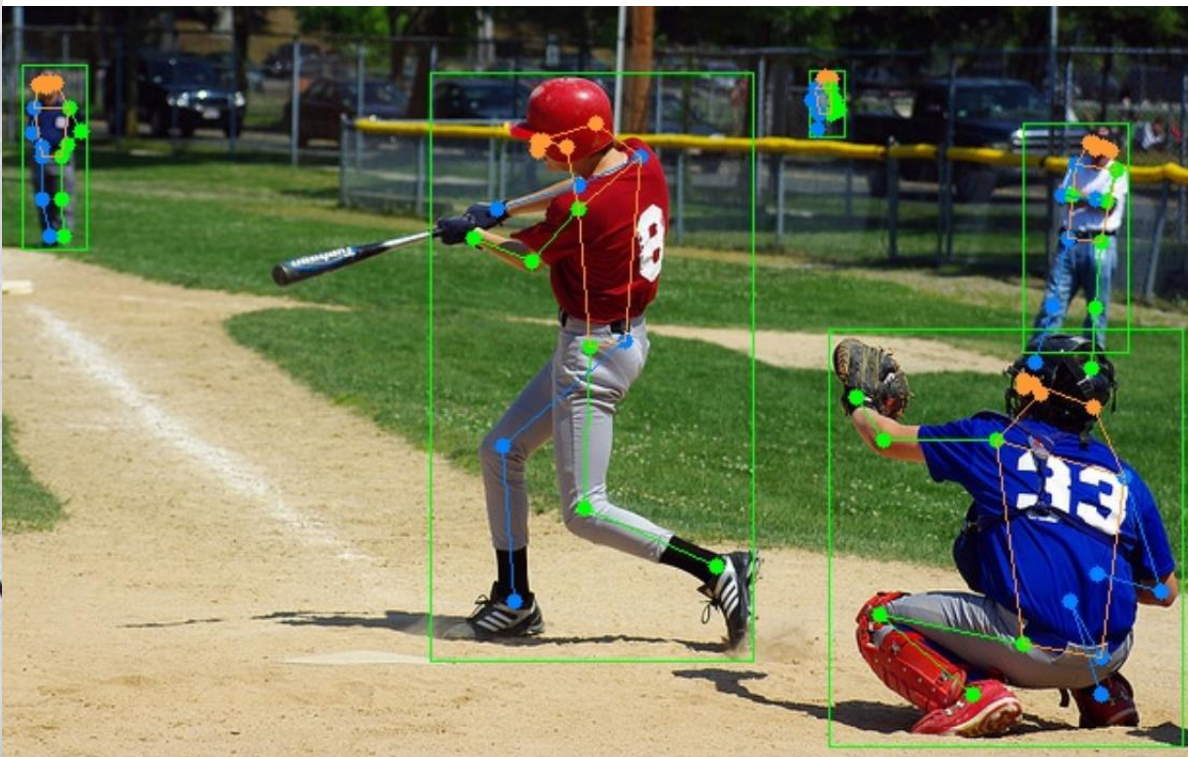
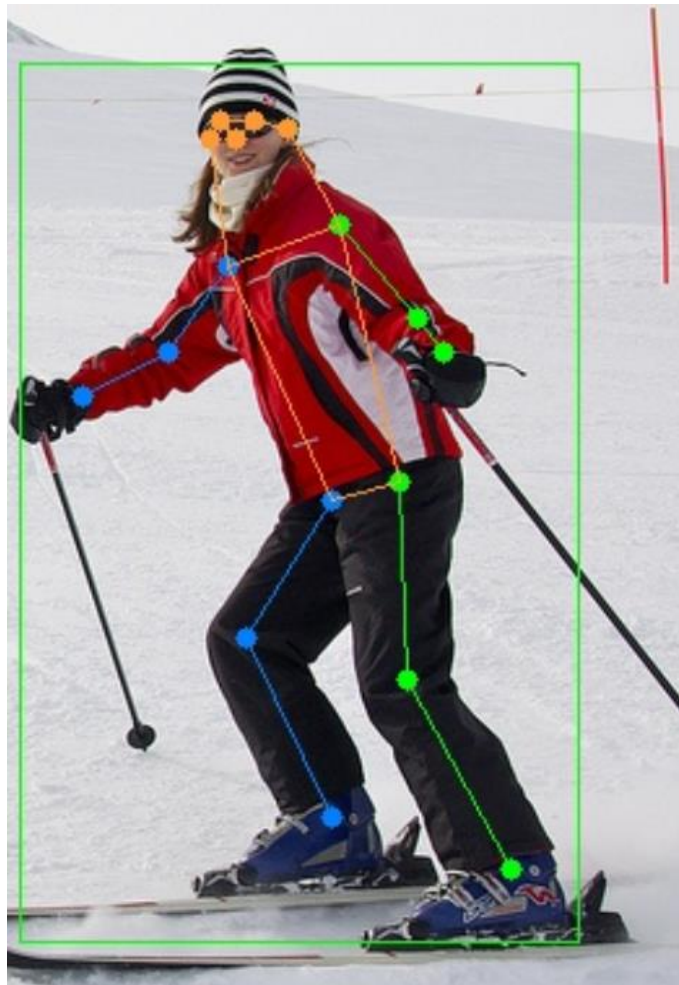
Frame: 1887

People: 1

(OpenPose を使用)



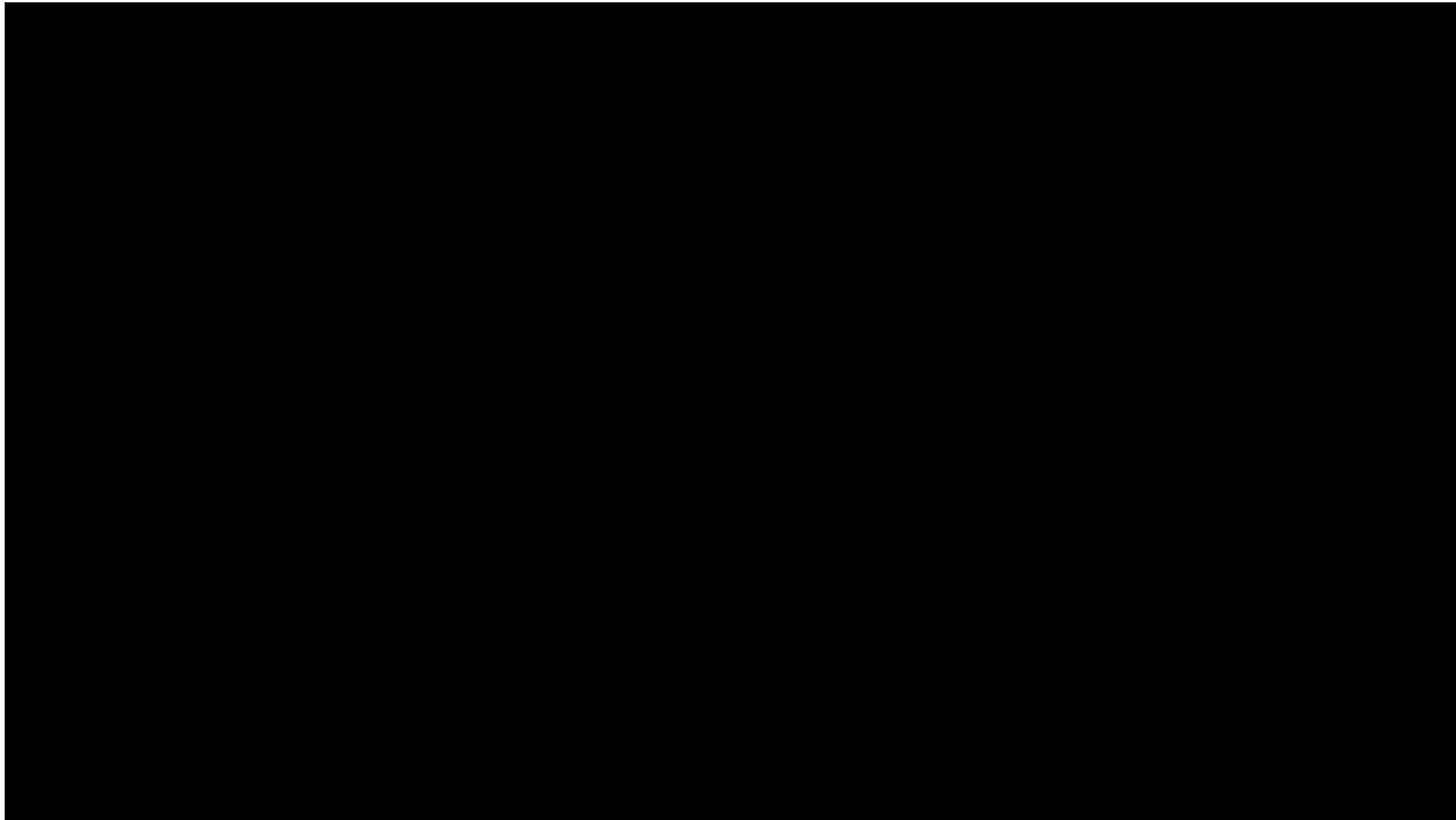
(Detectron2 の DensePose を使用)



ビデオの例



ビデオの例



姿勢推定（人やオブジェクトの動きを読み解く）

- **人間の全身**：人間の動きや行動。その人の活動や意図
- **手や足の指**：手指や足指の動き。細やかな指示や意図
- **頭部**：視線や注視の向き。意図や関心の対象
- **オブジェクト**：オブジェクトのふるまい。オブジェクトと人の相互作用。

姿勢推定の多様な応用分野

- **監視とセキュリティ**：公共の場で、人々の行動を監視。危険や異常を検出
- **エンターテインメント**：キャラクター・アニメーションの元データを作成
- **ヒューマンコンピュータインタラクション**：人間とコンピュータが相互作用するインタラクション
- **スポーツ分析**：選手の動きを分析。怪我を予防
- **医療、リハビリ**：身体の動きを評価。リハビリ効果の確認
- **頭部追跡の応用**：VR/AR、運転者の集中の確認など
- **オブジェクトの向き**の応用：ロボット、倉庫の荷物、自動運転車など

姿勢推定の用途①

- 状況把握，行動予測



姿勢推定の用途②

- 行動検知



姿勢推定の用途③

- 危険の察知, 救援・救護



姿勢推定の用途④

- 監視, 防犯



姿勢推定の用途⑤

- トレーニング, フィットネス, エクササイズ



姿勢推定の用途⑥

- アバター（バーチャル・キャラクター）の操作



GAN（敵対性生成ネットワーク）のニュース 

- 実在しない人間の顔画像を生成
tf-GAN, https://docs.google.com/presentation/d/1OpcYLBVpUFIL-wwPHu_CyKjXqXD0oRwBoGP2peScrSA/edit#slide=id.g4551faa5ed_0_208 より



- 色分け図や線画をリアルに変換
Video-to-Video Synthesis, <https://www.youtube.com/watch?v=S1OwOd-war8> より



23

アバター（バーチャル・キャラクター）

姿勢推定の用途⑦

- 集団行動の観察，分析，監視



(b)

Fig. 1 Examples of pose estimation results. (a) Single person pose estimation results from Ref. [5]. (b) Multi-person pose estimation results from Ref. [3].

姿勢推定の用途⑧

- 姿勢推定の結果を用いて，画像から人物を切り出す



ここまでのまとめ

姿勢推定は、人間の全身（人体）、頭部、およびその他のオブジェクトの**位置と方向**を推定する技術

要素

- **人間の全身**：人間の動きや行動。その人の活動や意図。
- **手や足の指**：手指や足指の動き。細やかな指示や意図。
- **頭部**：視線や注視の向き。意図や関心の対象。
- **オブジェクト**：オブジェクトのふるまい。オブジェクトと人の相互作用。

応用分野

- エンターテインメント、セキュリティ、ヒューマンコンピュータインタラクションなどさまざま

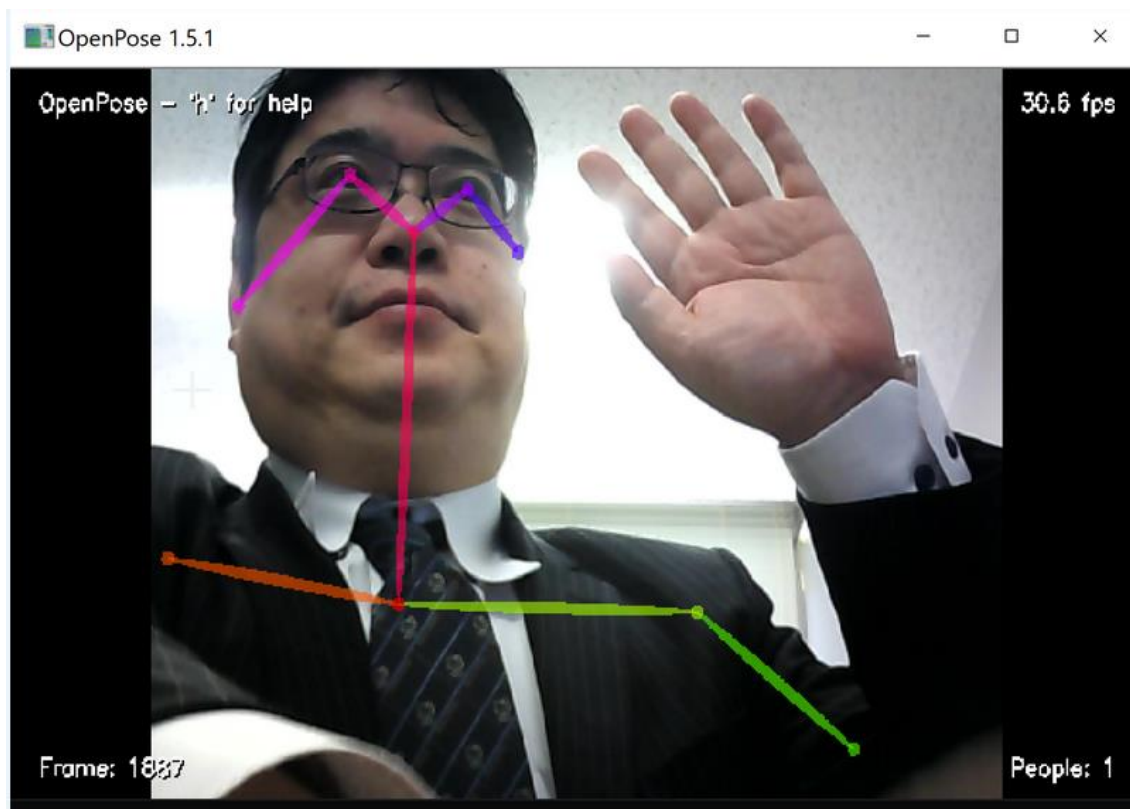


8-2. 人体の姿勢推定

人体の姿勢推定とキーポイント

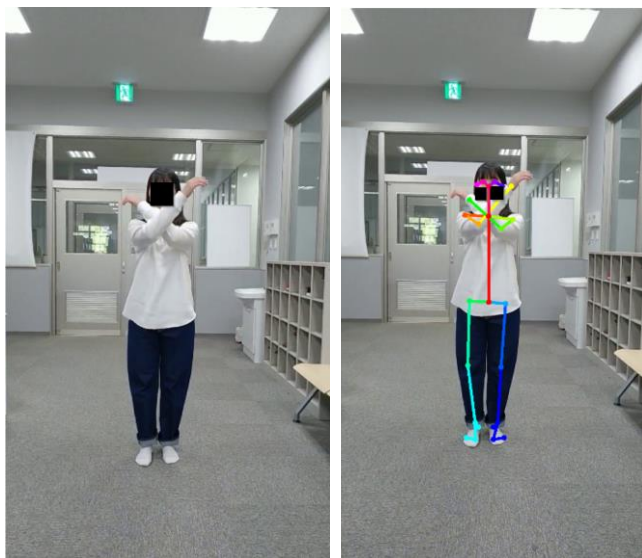
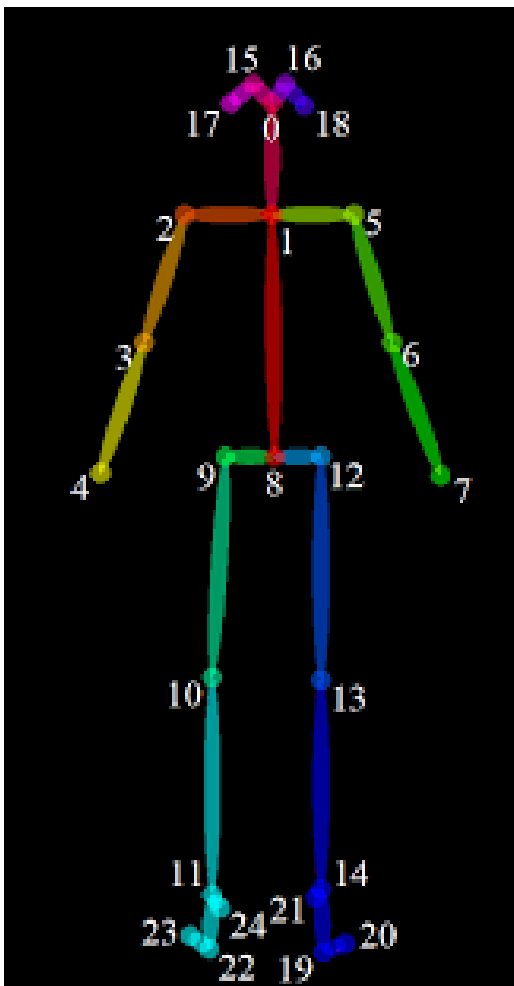


1. **キーポイントの検出**：姿勢推定では、人体の特定の「**キーポイント**」（特徴となるポイント）を検出し、**位置を特定**
2. **姿勢推定**：位置が特定された**キーポイントを結びつける**ことで、人間の全身の姿勢を推定



キーポイントとボーン

ボーン = 骨格, キーポイント = 特徴となるポイント



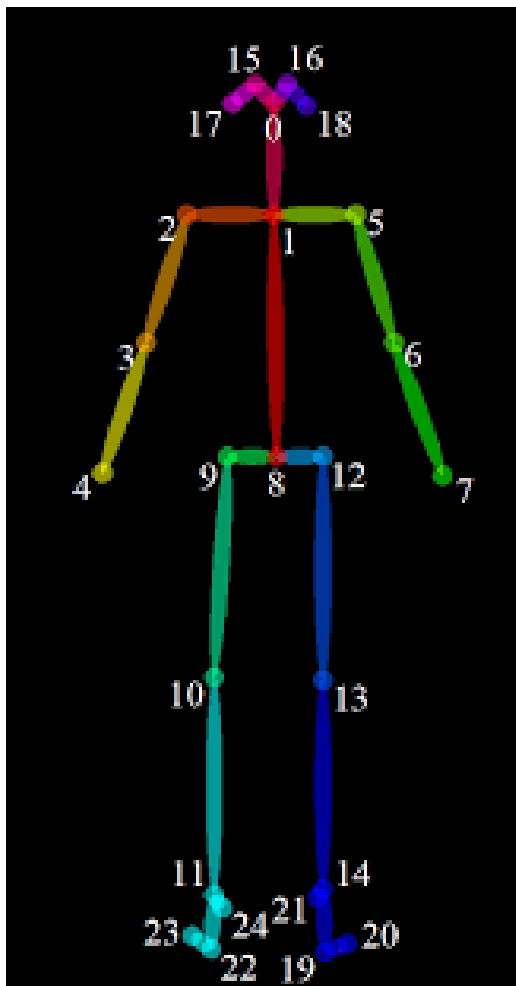
画像で表示した
ボーン情報

```
“version”: 1.3,  
“people”: [  
  {  
    “person_id”: [ 385.359,  
                  1008.49,  
                  -1 ],  
    “pose_keypoints_2d”:  
      363.314, 348.747, 0.868796,  
      432.965, 421.999, 0.967798,  
      0.979106, 0.967798,  
      363.37, 377.903, 0.93886,  
      506.265, 422.022, 0.956557,  
      304.656, 0.956557,  
      502.51, 333.997, 0.909367,  
      0.909367, 429.439,  
      337.621, 524.568, 0.802341,  
      0.840333, 396.402, 0.840333,  
      418.283, 433.016, 0.807274,  
      433.005, 0.594963, 422.06, 374.436,  
      513.559, 0.882184, 1048.78,  
      0.882184, 392.689, 0.816693,  
      535.576, 399.931, 0.725604,  
      0.725604, 315.785, 1045.14,  
      465.92, 0.858681,  
      0.342554, 374.45,  
      363.233, 1019.45,  
      715.201, 0.848111, 0.720292,  
      315.755, 315.755, 330.279,  
      711.596, 1048.79, 0.880071,  
      0.801404, 312.162, 0.880071,  
      869.036, 0.876103, 308.235,  
      0.876103, 330.276, 1044.99,  
      1008.47, 1008.47, 0.892973,  
      0.892966, 403.621, 337.594,  
      718.853, 1019.47, 0.792035,  
      0.792035, 392.782, 872.737,  
      0.901708, 0.901708, 0.761588
```

キーポイントのデータファイル
(特徴点の x, y 座標と
個々の信頼度のデータ)

キーポイントとボーン

ボーン = 骨格, キーポイント = 特徴となるポイント



OpenPose では、25個のキーポイントが定まっている

(OpenPose は、人体の姿勢推定の一手法)

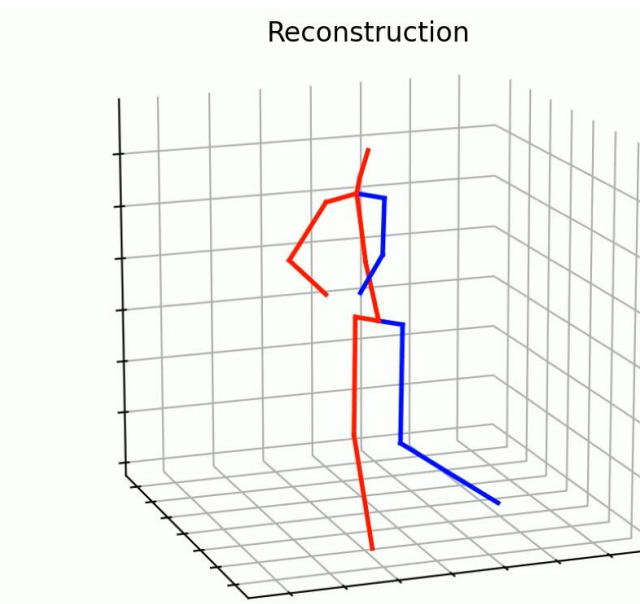
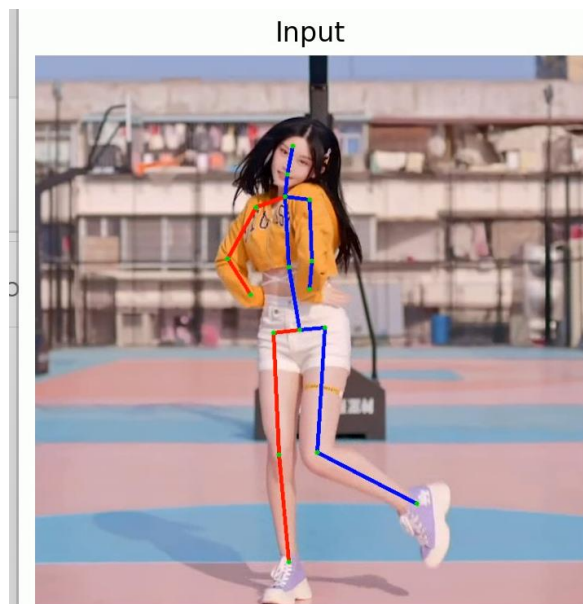
0:鼻 1:首 2:右肩 3:右ひじ 4:右手首
5:左肩 6:左ひじ 7:左手首 8:真ん中腰
9:右腰 10:右ひざ 11:右足首 12:左腰
13:左ひざ 14:左足首 15:右目 16:左目
17:右耳 18:左耳 19:左足親指 20:左足小指
21:左かかと 22:右足親指 23:右足小指
24:右かかと

2次元と3次元の姿勢推定

姿勢推定は**2次元（平面的な推定）**と**3次元（空間的な推定）**の2種類

- 2次元: 2次元的な（平面の）姿勢を推定
- 3次元: **空間的な深さ情報（奥行き）を含め**、3次元的な姿勢を推定

3次元の姿勢推定は、より詳細な情報を提供



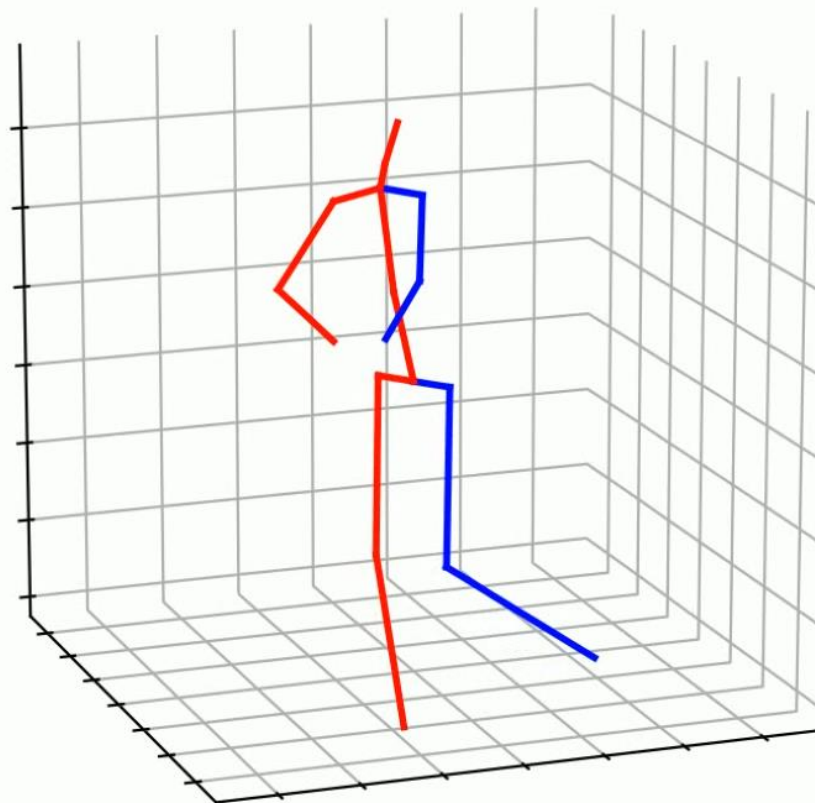
3次元姿勢推定

3次元の姿勢推定

Input

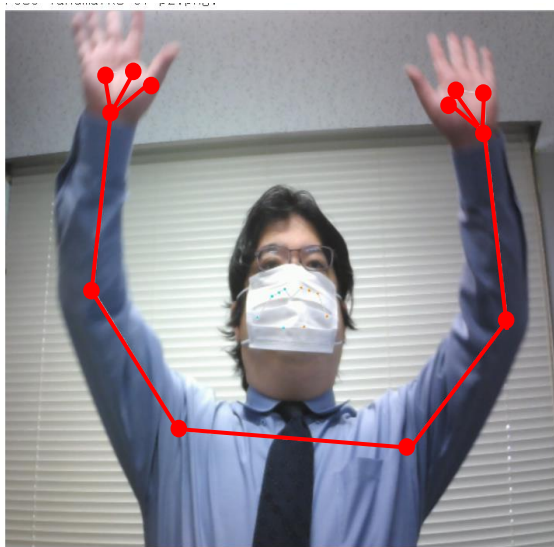


Reconstruction

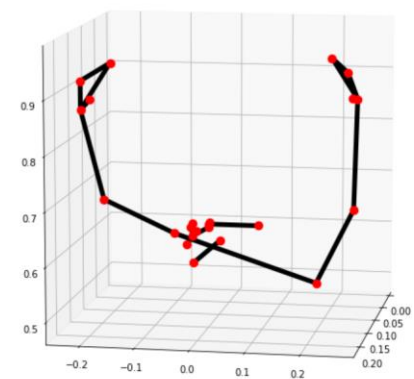


- **3次元の姿勢推定は、空間的な深さ情報（奥行き）を含め、3次元的な姿勢を推定**
- 人体表面のポリゴンメッシュを付けたり、アクションの分類を行うなどの応用

3次元の姿勢推定の例



キーポイントとボーン



キーポイントの位置情報は3次元（縦，横，奥行き）

ここまでのまとめ

人体の姿勢推定とキーポイント

- **姿勢推定**では、人体の特定の「**キーポイント**」（特徴となるポイント）を検出し、**位置を特定**
- **姿勢推定**：位置が特定されたキーポイントを結びつけることで、人間の全身の姿勢を推定
- OpenPose：人体の姿勢推定の一手法。25個のキーポイント（鼻、首、右肩、右ひじなど）

2次元と3次元の姿勢推定

- **2次元の姿勢推定**：2次元的な（平面の）姿勢を推定
- **3次元の姿勢推定**：空間的な深さ情報（奥行き）を含む、より詳細な姿勢推定

8-3. ディープラーニング による姿勢推定の仕組み

姿勢推定とディープラーニング



ディープラーニングの利用により、姿勢推定の**精度が向上**。

以下のような課題が部分的に解決

- 多人数の対応
- 人体の向きや姿勢のバリエーション
- 個々の人体の個性
- 部分遮蔽
- 照明のバリエーション（夜間、影）
- 衣服のバリエーション

さまざまな状況において、部分隠蔽があっても、姿勢推定ができるようになってきた

姿勢推定とディープラーニング

ディープラーニングは万能ではなく、次のような問題がある

- **訓練データのバリエーションの不足**

姿勢推定のために「制限のある実験環境」で撮影されたデータを使う場合、**訓練データのバリエーションが限られ**、その範囲でしか姿勢推定できない

この問題の解決のため、自己教師あり学習などのアプローチが行われている

ディープラーニングによる姿勢推定

1. キーポイントの検出

次の手順で、複数人体のキーポイント（特徴となるポイント）を識別

- ① **部位の位置推定**: 鼻、首、右肩、右ひじなどの**部位**を検出。
- ② **同一人物のキーポイント特定**: ステップ①の結果をもとに、**同一人物の部位**を集めて、最終的なキーポイントを得る

2. 姿勢推定

位置が特定された**キーポイントを結びつける**ことで、人間の全身の姿勢を推定

キーポイントの検出

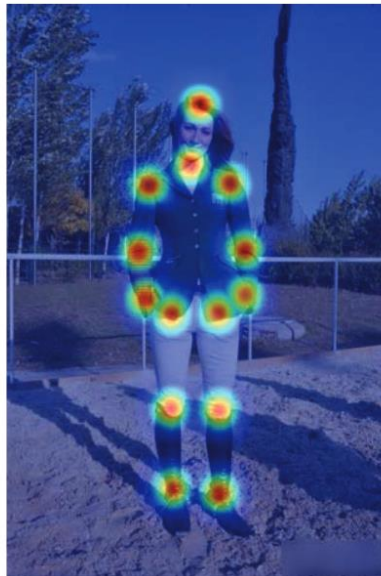
元画像

①部位の
位置推定

②同一人物の
キーポイント特定



(a)



(b)



(c)

Fig. 6 An example of heatmap-based single-person pipeline with heatmap. (a) Original image, (b) heatmap generated by estimator, and (c) detection result.

部位の位置推定を行う Hourglass

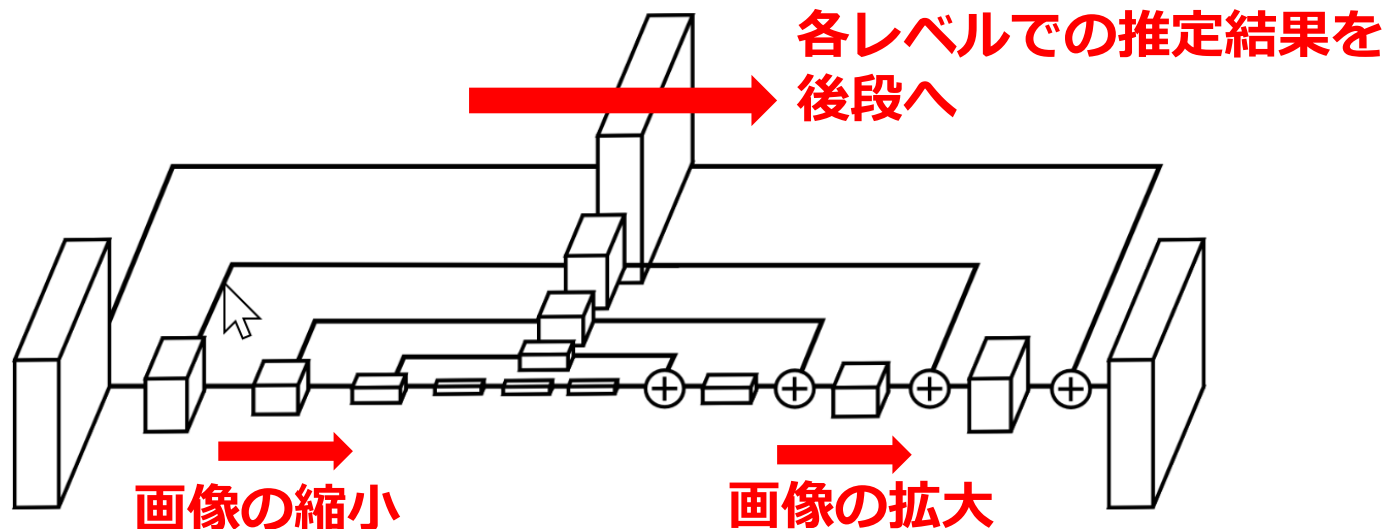


Fig. 3. An illustration of a single “hourglass” module. Each box in the figure corresponds to a residual module as seen in Figure 4. The number of features is consistent across the whole hourglass.

部位の位置推定

部位の大きさはさまざま ⇒ マルチスケールの画像を使う

- 画像を複数レベルに縮小し、**各レベルで部位**を検出
- 元の大きさに戻しながら、検出結果を重ね合わせ

部位の位置推定を行う Hourglass

- 画像の縮小： 畳み込み層とプーリング層で可能
- 画像の拡大： 逆畳み込み層あるいはアップサンプリング層で可能

Hourglass は、部位の位置推定に特化した畳み込みニューラルネットワーク

3次元の人体姿勢推定の技術

- MotionAGFormer (2023年)
- MHFormer (2022年)
- MotionBERT (2022年)
- DiffPose (2022年)

ディープラーニングによる3次元の姿勢推定

1. キーポイントの検出

次の手順で、複数人体のキーポイント（特徴となるポイント）を識別

- ① **部位の位置推定**: 鼻、首、右肩、右ひじなどの**部位**を検出。
- ② **同一人物のキーポイント特定**: ステップ①の結果をもとに、**同一人物の部位**を集めて、最終的なキーポイントを得る
- ③ ②の結果の**キーポイント**に対して、**空間的な深さ（奥行）**を求め、**3次元の姿勢推定**を実行 = 追加処理

2. 姿勢推定

位置が特定された**キーポイント**を結びつけることで、人間の全身の姿勢を推定

ここまでのまとめ

人体の姿勢推定の精度向上

多人数対応、姿勢のバリエーション、部分遮蔽、照明・衣服のバリエーションへの対応。

処理手順

- 部位の位置推定
- 同一人物のキーポイント特定
- キーポイントに対して深さ（奥行）を求める（3次元の姿勢推定を行う場合の追加処理）。

ディープラーニングによる姿勢推定の現在の課題

訓練データのバリエーション不足、自己教師あり学習のアプローチ。



演習

人体の3次元姿勢推定

1. 使用するページ:

https://colab.research.google.com/drive/1QRLqEAePmgS41vOKQwf87G_Ss_BLhPYs?usp=sharing#scrollTo=nSbrUG6xxz7r

(LART の研究者によるデモページ)

2. 目的

LART の作者によるデモページにより、
3次元の姿勢推定の理解を深める

3. 各自で行うこと

各自で説明、ソースコード、実行結果を確認する。



クリックして
展開

> Visualize the reconstructions

▶ 41 個のセルが非表示

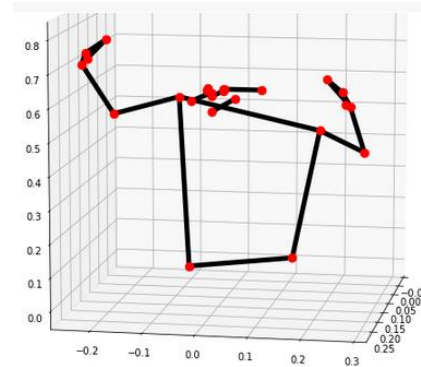
人体の3次元姿勢推定

1. 使用するページ:

https://colab.research.google.com/drive/13nOMSWODzx_LjN9XEG99jtvvgMACl4m9V

2. 目的

姿勢推定、キーポイント、ボーンについて理解を深める



3. 各自で行うこと

各自で説明、ソースコード、実行結果を確認する。

全体まとめ①

姿勢推定は、人間の全身（人体）、頭部、およびその他のオブジェクトの**位置と方向**を推定する技術

人体の姿勢推定とキーポイント

- **姿勢推定**では、人体の特定の「**キーポイント**」（特徴となるポイント）を検出し、**位置を特定**
- **姿勢推定**：位置が特定されたキーポイントを結びつけることで、人間の全身の姿勢を推定

2次元と3次元の姿勢推定

- **2次元の姿勢推定**：2次元的な（平面の）姿勢を推定
- **3次元の姿勢推定**：空間的な深さ情報（奥行き）を含む、より詳細な姿勢推定

全体まとめ②

ディープラーニングによる人体の姿勢推定

人体の姿勢推定の精度向上

多人数対応、姿勢のバリエーション、部分遮蔽、照明・衣服のバリエーションへの対応。

処理手順

- 部位の位置推定
- 同一人物のキーポイント特定
- キーポイントに対して深さ（奥行き）を求める（3次元の姿勢推定を行う場合の追加処理）

授業から得られる満足感

- キーポイントとボーンの概念、2D/3D姿勢推定の違い、ディープラーニングの進展と課題の理解向上
- 姿勢推定のアルゴリズムを学び、Pythonプログラムも確認。最新の研究トピックも知る。
- 監視・セキュリティ等での活用事例を知り、他技術との組み合わせによる新たな可能性を考察できるようになる
- 今後の自主的な学習やプロジェクトに活用できる知識とスキル向上