

# 卒業論文概要書

Summary of Bachelor's Thesis

(2018/1/30)

学科名 Department	情報通信	氏名 Name	稲田健太郎	指導 教員 Advisor	渡辺 裕 印 Seal
研究指導 Research guidance	オーディオビジュアル情報処理研究	学籍番号 Student ID number	1w142044-6 CD		
研究題目 Title	OpenPose を用いた複数人のダンスの一致度評価 Coincidence evaluation of multiple people's dance using OpenPose				

## 1.まえがき

モーションキャプチャやセンサを用いた動作解析では、複数台のカメラやマーカ・センサの準備が必要になる。複数台のカメラやセンサの前でしか動作解析ができない欠点がある。しかし、OpenPose はすでに撮影された動画で解析が行えるため、多人数の同時動作解析に適用できる。また、2012 年度からそれまで選択であった中学校での体育授業のダンス科目が必修化された[1]。ダンスがより身近なものとなり、複数人で踊る様子を撮影し公開する人も少なくない。そこで、特別な機具を持たない人が練習で利用できるダンスの評価指標が望まれている。本研究では、OpenPose の入力として複数人が同様のダンスをしている定点カメラ動画を利用する。ダンス動作のタイミングの一致度の評価をする手法を提案する。

## 2.OpenPose

OpenPose は、単一画像から複数の人間の体や顔のキーポイントをリアルタイムに検出することができる機械学習型動作解析処理である。Convolution Neural Network を用いて、画像に映る人物の肩・肘・目など 18 点の位置推定を行う[2]。動画を入力すると、全フレームに対して人物の姿勢推定を行う。各キーポイントにマークを付け、マーク間を線で結び人体モデルを表示する。また、フレームごとに映るすべての人物の各点座標のデータセットを同一ファイルで取得できる。

## 3.評価手法

OpenPose から得られた各キーポイントの座標のフレームに対する座標の変化の極値を求める。各人物の座標のピークから動作の一致度を評価する 2 つの手法を提案する。動作のタイミングのずれの許容フレームを  $K$  フレームとする。動画内の人物の過半数が  $K$  フレーム間にピークを持つときを正解の動作とする。手法 1 は、 $K$  フレーム間に過半数の人物がピークを持つ場合に動作の一致度の評価を行う。 $K$  フレーム間にピークを持つ場合を正解とし、ピークを持たない場合を不正解とする。これにより、各人物の各キーポイントでの正解数と不正解数を求める。動画内の各人物の正解数と不正解数にばらつきがある場合、全員の動作の一致度が低いと考えられる。手法 2 は、手法 1 の評価に加えて、 $K$  フレーム間にピークを持つ人物が過半数に満たない場合も評価を行う。このときの過半数に満たない人物がピークを持つ回数は、評価対象者のダンスの習熟度に依存する。

## 4.実験結果

5 人のアイドルグループがダンスする動画のうち、5 人は移動せず同じ順番で並び同じダンスをする一部分を使用した。タイミングのずれの許容フレーム数である  $K$  を 3 と設定した。5 人の各キーポイントでの正解数と不正解数から、正解率を求めた。その結果から、各人物の動作の一致度を数値に評価できた。また、各人物の正解率が高くないため、全員のダンス熟練度は高くないと考えられる。また、3 フレーム間にピークを持つ人物が過半数に満たない回数を求めた。この回数から、各人物の見直すべき回数がわかる。

## 5.まとめ

本研究では、複数人が同様のダンスを行う動画内の各人物の動作のタイミングの一致度を OpenPose により数値で評価する研究を行った。動画内の人物の過半数が  $K$  フレーム間にピークを持つときを正解の動作とすることで、各人物が正解の動作をした回数と正解の動作をしなかった回数から、正解率を求めた。各人物の一致度を数値で評価できることを示した。また、各人物の正解数・不正解数と  $K$  フレーム間に動画内の過半数に満たない人物がピークを持つ回数から、評価対象者のダンスの習熟度の評価も行った。しかし、実際に動画を見た人による主観評価と数値の関係が不透明なため、動画の主観評価を実験する必要がある。また、習熟度と各人物の正解率を評価する方法を検討する必要がある。

## 参考文献

- [1] 文部科学省, 新学習指導要領に基づく中学校向け「ダンス」リーフレット, [http://www.mext.go.jp/a\\_menu/sports/jyujitsu/1306098.htm](http://www.mext.go.jp/a_menu/sports/jyujitsu/1306098.htm), May, 2011
- [2] Z.Cao,T.Simon,S-E Wei,Y.Sheikh,“Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”, <https://arxiv.org/abs/1611.08050>,2016

2017 年度 卒業論文

# OpenPose を用いた複数ダンサーの動作同期評価

## Evaluation of Action Synchronization for Multiple Dancers Using OpenPose

指導教員 渡辺 裕 教授

早稲田大学 基幹理工学部

情報通信学科

1w142044-6

稲田 健太郎

# 目次

第 1 章	序論 .....	1
1.1	研究の背景 .....	1
1.2	関連研究 .....	1
1.3	研究の目的 .....	1
1.4	論文の構成 .....	2
第 2 章	OpenPose について .....	3
2.1	まえがき .....	3
2.2	OpenPose の概要 .....	3
2.3	利点と課題 .....	4
第 3 章	提案手法 .....	6
3.1	まえがき .....	6
3.2	データ整形 .....	6
3.3	角度による評価 .....	6
3.4	座標による評価 .....	7
3.4.1	ピーク検出 .....	7
3.4.2	採点手法 .....	8
第 4 章	実験結果と考察 .....	10
4.1	複数人数のダンス映像 .....	10
4.2	実験結果 .....	10
4.2.1	手法 1 .....	12
4.2.2	手法 2 .....	13
4.3	考察 .....	15
第 5 章	結論 .....	17
5.1	まとめ .....	17
5.2	課題 .....	17
	謝辞 .....	18
	参考文献 .....	19
	図一覧 .....	20
	表一覧 .....	20
	研究業績 .....	22

# 第1章 序論

## 1.1 研究の背景

前期中等教育の体育授業では、それまで選択性だったダンス科目が2012年度から必修化された[1]。また、身近なコンテンツであるアイドルグループやアーティストが簡単なダンスを踊りながら歌うことが多くなっている。そのダンスが流行となり、複数人数で踊る様子をスマートフォンなどで動画に撮り「踊ってみた動画」と名付けてYouTubeなどの動画サイトにアップロードする人が増加している。それらの動画の視聴者は、動きのキレやかわいらしさなどといった様々な尺度でダンスの評価する。以上のことから、若い世代にとってダンスとは、テレビ等で見えるものから実際に自分たちが踊るものおよび他人に評価されるものに変化している。

複数人数のダンスを動画に撮り、それを元に各人物の採点をすることができれば、ダンスの練習時に一つの指標となり得る。

## 1.2 関連研究

複数人動画像からの異常検出の関連研究として、南里ら[2]は動画像から立体高次局所自己相関特徴を用いた異常動作検出の手法を提案している。この手法では、背景と検出対象者を分離するために二値化した時間差分画像を用いる。まず、「歩く」を通常の動作として、特徴ベクトルを学習させる。もし「転ぶ」「走る」という動作をする人物がいる場合、学習したベクトルとは異なるベクトルが加わることを利用して、異常を検出している。しかし、複数人数が同一画像に映っている場合に異常者がいるかどうかのみを判断しており、その人物の検出までは行っていない。また、「歩く」と「転ぶ」という大きく動作に差がある場合のみであり、ダンスなど細かい動作には対応していない。

## 1.3 研究の目的

動画から複数人数の動作を採点することができれば、客観的に動作を見直しダンスの練習に利用できる。つまり、指導者がいない素人のみの集団でも、複数人数でダンスを揃えることが容易になる。そこで、本研究では各人物の動作のタイミングが周りの人物とどれだけ揃っているかを数値で評価することを目的とする。

## 1.4 論文の構成

以下に本章以降の構成を示す.

第 1 章は本章であり, 本論文の研究の背景および目的について述べている. また, 関連研究について述べている.

第 2 章では, **OpenPose** の概要および本研究における **OpenPose** の利点と課題について説明する.

第 3 章では, 動画内に映る人物の採点方法を提案する.

第 4 章では, 第 3 章で述べた提案手法における実験および結果について示す.

第 5 章では, 本研究のまとめと今後の課題を示す.

## 第2章 OpenPose について

### 2.1 まえがき

本章では、本論文で用いる OpenPose について述べる。OpenPose は Zhe ら[3]により提案された機械学習ベースの画像内の人物の身体のパーツ位置推定の手法であり、GitHub から無償公開されている人物の姿勢推定解析処理である。

### 2.2 OpenPose の概要

OpenPose とは、単一画像から複数の人間の身体や顔のキーポイントをリアルタイムに検出することができる機械学習型動作解析処理である[3]。Convolution Neural Network を用いて、画像に映る人物の肩や肘など 18 点の位置推定を行う。画像を入力すると、画像から検出した各点にマークとマーク間を線で結び人体モデルを表示した出力画像を得られる。また、各点の座標データを JSON, XML, YML 形式の出力ファイルとして得られる。なお、検出できないキーポイントの座標データは、 $x$ 座標と $y$ 座標ともに 0 となる。図 2.1 を入力画像とした場合、図 2.2 のような出力画像になる。



図 2.1 入力画像[4]



図 2.2 出力画像[4]

また、OpenPose は画像だけでなく動画に対しても利用可能である。動画を入力した場合、動画の各フレームに対して人物の姿勢推定を行い、座標データをフレームごとに別ファイルで出力される。

### 2.3 利点と課題

人物の動作解析では OpenPose 以外にモーションキャプチャやレーザーセンサを用いた手法がある。モーションキャプチャとは、マーカを取り付けた計測対象を複数台のカメラで撮影し、人物の動作を計測するアプローチである。スポーツ分野での人物の動作データの収集だけでなく、映画やゲームにおける CG で作成されたキャラクターの動作の再現にも使用されている。レーザーセンサを用いた動作解析は、対象物に向けて当てたレーザーが反射して戻るまでの時間から距離を測定し、人物の動作を 3D で捕らえることができる。モーションキャプチャとは違いマーカを付ける必要がないため、人物の動作を制限しない。これらの方法では複数台のカメラ・マーカまたはレーザーセンサが必要となる。また、解析が撮影時にカメラやセンサの前でしか行えない。一方で、OpenPose はウェブカメラで撮影された動画や録画されている動画に対してキーポイントの位置推定が行える。

このため、特別な機器が必要ない上に撮影場所が自由である。よって、他の動作解析方法より多くの人物の解析に利用できる OpenPose による動作解析が、本研究に適している。

しかし、画像内で人物同士や人物と物が重なると誤検出が多くなる。すべてのキーポイントを検出できない場合や、ある検出人物と重なる人物の身体の一部を検出人物の身体の一部だと誤って検出してしまう場合がある。人物と物が重なりすべてのキーポイントを検出できない例を図 2.3 に示す。

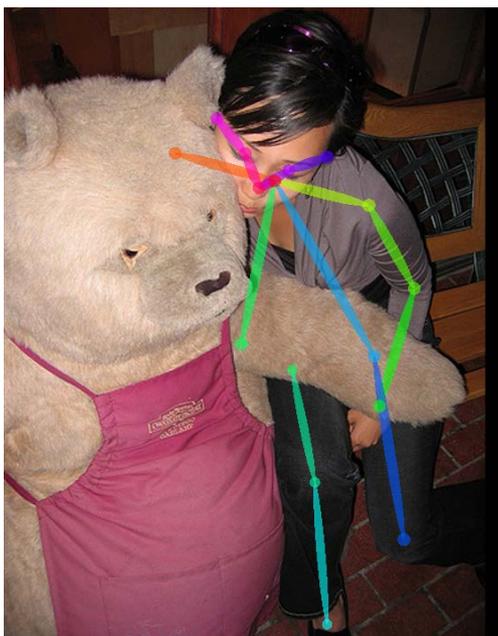


図 2.3 誤検出の例[4]

## 第3章 提案手法

### 3.1 まえがき

本章では, OpenPose を用いて, 複数人が同様のダンスをしている定点カメラ動画における動作のタイミング一致度の評価をする手法を提案する. 3.2-3.4 において, 処理の詳細について述べる. 本研究では, OpenPose により得られた 18 点の座標データのうち, 首・両肩・両肘・両手首・両腰・両膝・両足首の 13 点を処理対象とする.

### 3.2 データ整形

得られた複数人の座標データから人物を推定する. 動画内で人物の立ち位置が変わらない場合, 各人物の首の  $x$  座標に大きな変化がない. これにより, 首の  $x$  座標を 1 フレーム前の各人物の首の  $x$  座標と比較することで人物を推定できる. 1 フレーム前の各人物の首の  $x$  座標との差が最も小さい座標データをその人物になる.

また, 座標データが欠損している場合は, 前後フレームの座標データを参照する. 欠損フレームの直前フレームの座標から直後フレームの座標まで一定に変化するものとして, 欠損フレームの座標データを決定する.

### 3.3 角度による評価

ある動作をしているときの腕や脚の角度は, 人物の身長や腕の長さに関わらず一定となることが多いと考えられる. したがって, 時系列における角度の増減量により, 動作の速さが確認できる. また, 時系列における角度変化の極値によりある動作から次の動作に変わるタイミングが確認できる. 動作の速さとタイミングから複数人の一致度が評価できる. そこで, 3.2 で示したデータ整形で得られた座標から左右の首と肩と肘, 肩と肘と手首, 腰と膝と足首が成す 6 角を求め, その推移から一致度を評価する. 点  $A, B, C$  の座標が与えられたとき,  $\angle ABC$  を  $\theta^\circ$  とする.  $BA$  から  $BC$  に回転する方向が左回りのならば,  $\theta > 0$  と定義する. このとき,  $\theta$  の大きさは式(3.1)で与えられ, 回転方向は外積  $\vec{BA} \times \vec{BC}$  の符号と等しくなる.

$$\theta = \cos^{-1} \left( \frac{\vec{BA} \cdot \vec{BC}}{|\vec{BA}| |\vec{BC}|} \right) \quad (0 < \theta < 180) \quad (3.1)$$

この手法では身体の向きの見え方が動画内の立ち位置により大きく変わる. 評価対象全

員がカメラに対して正面を向いている時は、動画内での身体の向きの見え方が等しい。しかし、カメラに対して少しでも左右に向くと腕や脚の角度に影響を及ぼすため、角度の極値が正確ではなく評価ができない。身体の向きを変える動作はダンスでは一般的であるため、本研究では角度による評価手法は有効ではないため、使用しなかった。

### 3.4 座標による評価

角度とは違い、身長や腕の長さは人物によって座標の増減量が異なる。そのため、座標の増減量から動作の評価はできない。しかし、複数人がある同じ動作をするとき、身長や腕の長さに関わらず同時に時系列における座標変化の極値が生じる。その極値はある動作から次の動作に変わる瞬間であるため、そのタイミングで動作の評価が可能となる。座標変化の極値のタイミングは、動画内の立ち位置と身体の向きによる影響がない。そこで本研究では、身体の向きが変わる動作にも有効である座標変化の極値による評価手法を採用する。

#### 3.4.1 ピーク検出

座標の増減から動作の開始と終了のタイミングを判断するためピークを求める。フレームに対する座標の変化の波形の中で、極大値と極小値を求める。13点のキーポイントの $x$ 座標と $y$ 座標の合計26個の要素に対して行う。あるピークに対して、直前のピークのフレームをピーク開始点とする。動画内の任意の人物とキーポイントに対して、 $n$ フレーム目の座標を $X_n (n = 1, 2, 3, \dots, N)$ と、 $m$ 個目のピークのフレームを $PF_m (m = 1, 2, 3, \dots, M)$ とする。式(3.3)を満たすとき $PF_m = n$ となる。また、フレームに対する座標変化のグラフの例を図3.1に示す。図3.1のプロットされた点はピークを示し、そのフレームを $PF_m$ となる。

$$(X_n - X_{n-1})(X_{n+1} - X_n) < 0 \quad (3.3)$$

ここで、 $m$ 個目と $m - 1$ 個目のピークの座標の差を $PH_m$ とした時、式(3.4)のようになる。

$$PH_m = \begin{cases} X_{PF_m} - X_1 & m = 1 \\ X_{PF_m} - X_{PF_{m-1}} & m \neq 1 \end{cases} \quad (3.4)$$

ピーク間の座標の差が小さいものは、意図された動作ではないと仮定する。それらを解析対象から除去するために、ピーク間の座標差にしきい値を設定する。動作が大きい肘と手首には、しきい値 $T_1$ を設定し、その他のキーポイントには $T_1$ より小さいしきい値 $T_2$ を設定した。

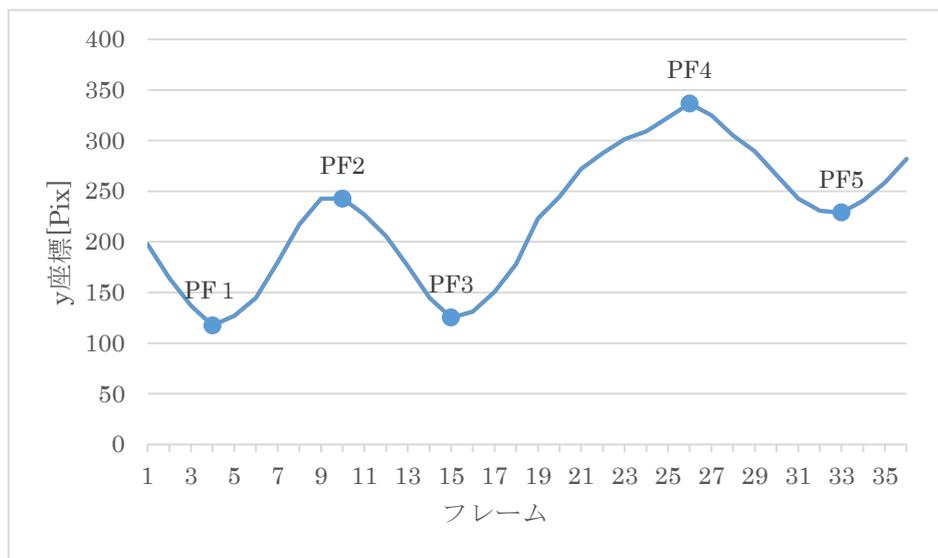


図 3.1 座標変化の例

### 3.4.2 採点手法

各人物の座標のピークから動作の一致度を評価する二つの手法を提案する。ここで、動作のタイミングのずれの許容フレーム数を $K$ フレームとする。本研究では、動画内の人物の過半数が $K$ フレーム間にピークを持つときを正解の動作とし、動画内の人物に対して相対的な評価した。各キーポイントの $x$ 座標と $y$ 座標で評価するだけでなく、肩・肘・手首の合計を腕部分、腰・膝・足首の合計を脚部分とし、それらでも評価する。動画内での人物の過半数が一致している動作を正解とすることにより、ダンスごとの動作の正解モデルの準備は不要である。

#### 3.4.2.1 手法 1

動画内の各人物の各キーポイントについて、 $K$ フレーム間に過半数の人物がピークを持つ場合に動作の一致度を評価する。このとき、 $K$ フレーム間にピークを持つ場合を正解とし、ピークを持たない場合を不正解とする。これにより、各人物の各キーポイントでの正解数と不正解数を求め、正解率を計算する。結果より、正解率が大きいほど、その人物が決められた動作をしていると考えられる。一方で、正解率が小さいほど、その人物が決められた動作をしていないと考えられる。

また、各人物の正解率の大きさにより、評価対象者のダンス習熟度を評価する。ダンス熟練者の集団の場合、全員が正確に決められた動作をしていると考えられる。つまり、全員の不正解数が少ないため、各人物の正解率が大きいと考えられる。しかし、ダンス初心者の集

団の場合、決められた動作をしていない人物や周りとのタイミングが一致しない人物がいると考えられる。このため、各人物の正解率が小さいと考えられる。

#### 3.4.2.2 手法 2

$K$ フレーム間にピークを持つ人物が過半数に満たない場合を評価する。このときの過半数に満たない人物がピークを持つ回数は、評価対象者のダンスの習熟度に依存する。ダンス熟練者の集団の場合は、全員が正確に踊れているため、この回数は小さいと考えられる。一方、ダンス初心者の集団の場合は、誤った動作をしていると考えられるため、この回数は大きいと考えられる。

## 第4章 実験結果と考察

### 4.1 複数人数のダンス映像

本研究では、5人のアイドルグループがダンスする動画のうち、5人は移動せず同じ順番で並び同じダンスをする一部分を使用した。この動画は23fpsの約9秒間で、全フレーム数は228である。人物を左からA,B,C,D,Eとする。この動画の第1フレームの人体モデルを図4.1に示す。

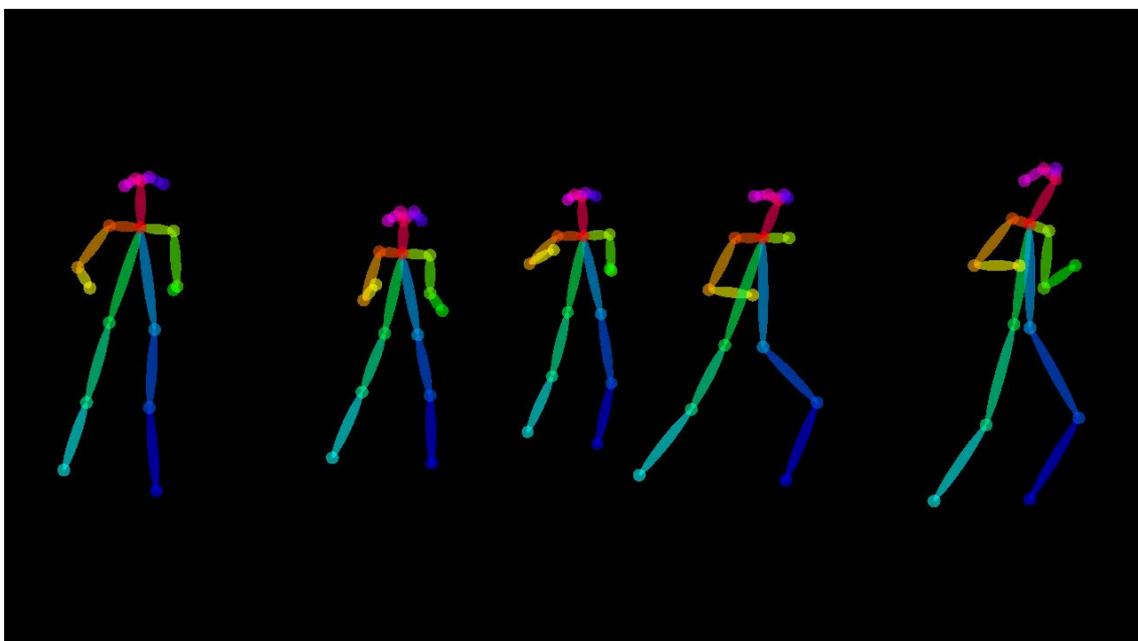


図 4.1 使用動画の第1フレーム

### 4.2 実験結果

実験では、実験条件として表4.1に示すパラメータの値を使用した。

表 4.1 実験条件

パラメータ	値
$T_1$	10
$T_2$	5
$K$	3

3.4.2 の二つの手法でそれぞれ評価した。動画全体での評価だけでは、一つの振り付けに対しての動作の一致度がわからない。そのため、第 115 フレームから第 150 フレームまでの手を大きく振る振り付けに対しても評価した。得られた座標データのうち、5 人の右手首の y 座標の時系列における変化の様子を図 4.1 に示す。また、第 115 フレームから第 150 フレームまでの 5 人の右手首の y 座標の時系列における変化の様子を図 4.2 に示す。

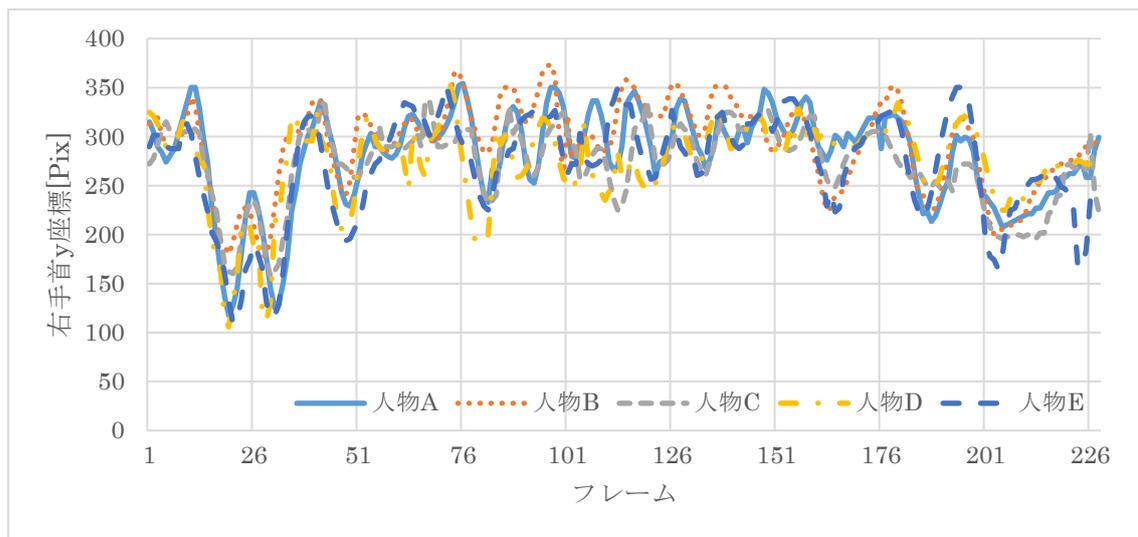


図 4.2 動画全体の 5 人の右手首 y 座標の変化

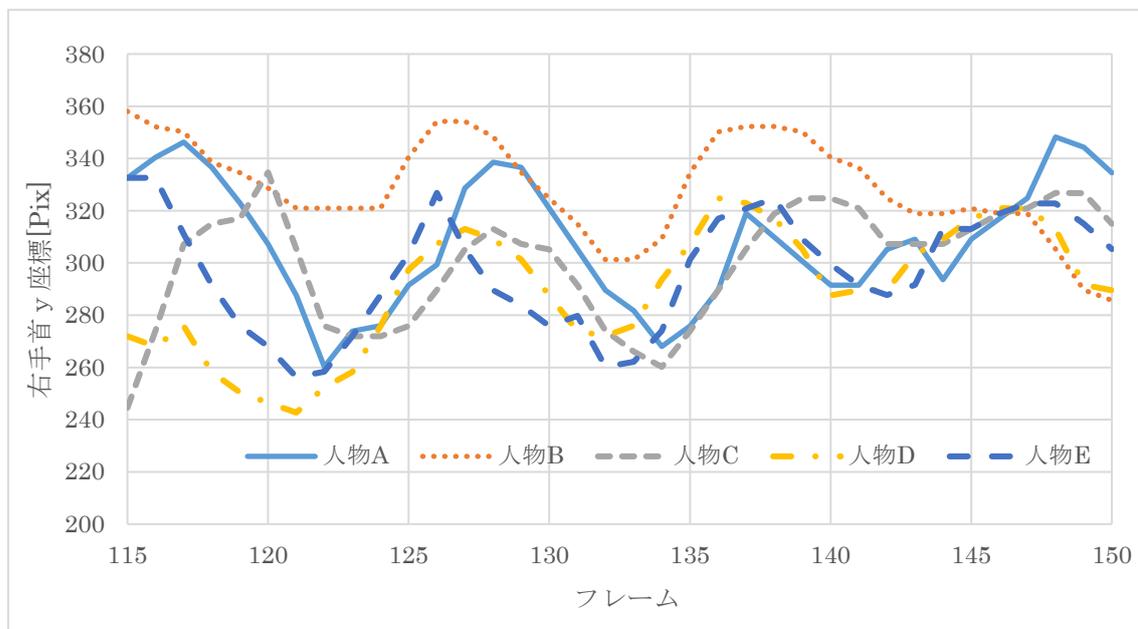


図 4.3 動画一部分の 5 人の右手首 y 座標の変化

#### 4.2.1 手法 1

3.4.2.1 で述べた手法 1 による実験結果を示す。動画全体での各キーポイントの $x$ 座標と $y$ 座標の合計の正解率を表 4.2 に、左右の腕部分と脚部分と全身の正解率を表 4.3 に示す。また、第 115 フレームから第 150 フレームまでの各キーポイントの合計の正解率を表 4.4 に、左右の腕部分と脚部分と全身の合計の正解率を表 4.5 に示す。

表 4.2 動画全体の各キーポイントの正解率

	人物 A	人物 B	人物 C	人物 D	人物 E
首	53.2%	83.0%	72.3%	59.6%	66.0%
右肩	70.9%	65.5%	65.5%	58.2%	61.8%
右肘	72.1%	58.8%	60.3%	60.3%	60.3%
右手首	59.0%	60.7%	60.7%	54.1%	62.3%
左肩	45.8%	77.1%	66.7%	54.2%	75.0%
左肘	69.6%	62.5%	39.3%	60.7%	48.2%
左手首	71.6%	64.2%	47.8%	50.7%	56.7%
右腰	51.0%	71.4%	69.4%	59.2%	59.2%
右膝	78.9%	75.4%	47.4%	52.6%	45.6%
右足首	83.3%	71.4%	31.0%	61.9%	47.6%
左腰	65.9%	68.2%	70.5%	65.9%	52.3%
左膝	53.7%	66.7%	59.3%	51.9%	61.1%
左足首	66.7%	69.2%	38.5%	66.7%	56.4%

表 4.3 動画全体の腕・脚部分と全身の正解率

	人物 A	人物 B	人物 C	人物 D	人物 E
右腕	67.4%	61.4%	62.0%	57.6%	61.4%
左腕	63.7%	67.3%	50.3%	55.0%	59.1%
右脚	70.9%	73.0%	50.0%	57.4%	50.7%
左脚	61.3%	67.9%	56.9%	60.6%	56.9%
全身	65.1%	68.1%	56.2%	57.6%	57.9%

表 4.4 動画一部分の各キーポイントの正解率

	人物 A	人物 B	人物 C	人物 D	人物 E
首	40.0%	90.0%	50.0%	70.0%	30.0%
右肩	88.9%	88.9%	66.7%	77.8%	22.2%
右肘	57.1%	64.3%	57.1%	64.3%	64.3%
右手首	46.2%	53.8%	53.8%	53.8%	76.9%
左肩	22.2%	88.9%	55.6%	55.6%	77.8%
左肘	33.3%	77.8%	33.3%	100.0%	66.7%
左手首	70.0%	50.0%	60.0%	80.0%	50.0%
右腰	44.4%	44.4%	55.6%	77.8%	77.8%
右膝	90.9%	81.8%	45.5%	36.4%	27.3%
右足首	66.7%	83.3%	16.7%	50.0%	33.3%
左腰	60.0%	70.0%	40.0%	70.0%	70.0%
左膝	70.0%	60.0%	60.0%	60.0%	50.0%
左足首	50.0%	66.7%	66.7%	66.7%	50.0%

表 4.5 動画一部分の腕・脚部分と全身の正解数と不正解数

	人物 A	人物 B	人物 C	人物 D	人物 E
右腕	61.1%	66.7%	58.3%	63.9%	58.3%
左腕	42.9%	71.4%	50.0%	78.6%	64.3%
右脚	69.2%	69.2%	42.3%	53.8%	46.2%
左脚	61.5%	65.4%	53.8%	65.4%	57.7%
全身	57.1%	69.8%	51.6%	65.9%	54.8%

#### 4.2.2 手法 2

3.4.2.2 で述べた手法 2 による実験結果を示す。3 フレーム間に 2 人がピークを持つ回数と 1 人しかピークを持たない回数を求めた。動画全体の各キーポイントのそれぞれの回数を表 4.6 に、左右の腕部分と脚部分と全身のそれぞれの回数を表 4.7 に示す。また、第 115 フレームから第 150 フレームまでの各キーポイントのそれぞれの回数を表 4.8 に、左右の腕部分と脚部分と全身のそれぞれの回数を表 4.9 に示す。

表 4.6 動画全体の各キーポイントのピークの人数ごとの回数

	人物 A		人物 B		人物 C		人物 D		人物 E	
	1 人	2 人	1 人	2 人	1 人	2 人	1 人	2 人	1 人	2 人
首	20	8	5	5	13	9	8	9	8	5
右肩	24	17	7	6	10	5	5	9	12	7
右肘	20	18	8	9	11	2	11	8	16	7
右手首	18	10	8	6	10	15	17	7	9	10
左肩	15	17	6	9	7	11	16	14	11	7
左肘	13	13	9	8	10	13	19	15	13	7
左手首	13	10	16	4	11	10	21	10	6	8
右腰	17	13	11	12	7	5	14	11	13	7
右膝	13	11	8	12	12	9	10	11	10	14
右足首	15	13	11	5	6	5	13	12	20	9
左腰	16	12	11	5	9	8	11	10	12	7
左膝	15	11	18	7	7	3	10	9	7	14
左足首	18	16	15	6	5	7	11	12	18	13

表 4.7 動画全体の腕・脚部分のピークの人数ごとの回数

	人物 A		人物 B		人物 C		人物 D		人物 E	
	1 人	2 人	1 人	2 人	1 人	2 人	1 人	2 人	1 人	2 人
右腕	62	45	23	21	31	22	33	24	37	24
左腕	41	40	31	21	28	34	56	39	30	22
右脚	45	37	30	29	25	19	37	34	43	30
左脚	49	39	44	18	21	18	32	31	37	34
全身	197	161	128	89	105	93	158	128	147	110

表 4.8 動画一部分の各キーポイントのピークの人数ごとの回数

	人物 A		人物 B		人物 C		人物 D		人物 E	
	1人	2人								
首	6	1	2	0	2	2	0	0	1	1
右肩	9	1	1	0	1	0	1	1	3	2
右肘	3	4	1	3	0	1	0	0	1	2
右手首	3	3	0	0	0	2	1	0	1	1
左肩	1	4	1	3	1	1	1	1	0	1
左肘	3	3	0	0	4	2	3	2	2	3
左手首	3	3	0	1	1	3	1	1	2	2
右腰	2	2	1	2	1	2	2	1	1	1
右膝	4	1	1	1	1	2	1	3	1	3
右足首	4	6	0	3	0	0	2	2	4	3
左腰	3	2	1	0	1	2	1	1	4	1
左膝	2	0	2	0	0	0	3	0	3	0
左足首	3	3	1	1	0	1	2	1	3	2

表 4.9 動画一部分の腕・脚部分のピークの人数ごとの回数

	人物 A		人物 B		人物 C		人物 D		人物 E	
	1人	2人								
右腕	15	8	2	3	1	3	2	1	5	5
左腕	7	10	1	4	6	6	5	4	4	6
右脚	10	9	2	6	2	4	5	6	6	7
左脚	8	5	4	1	1	3	6	2	10	3
全身	46	33	11	14	12	18	18	13	26	22

### 4.3 考察

表 4.2 から表 4.5 より、動画全体と動画一部分における各人物の動作の正解率がわかる。動画全体と動画一部分では、正解率に違いがある。これは、ダンスの一つの振り付けの合計が動画全体の正解率となるからである。動画一部分での正解率が大きい場合、その部分の動作は周りの人物と一致していることがわかる。一方で、動画一部分での正解率が

小さい場合、その部分の動作は周りの人物と一致していないため、練習が必要であることがわかる。よって、一つひとつの振り付けに注目して評価することで、各人物の練習が必要な部分が出る。また、各人物の正解率が60%程度であるため、5人はダンスが上手い集団ではないことが考えられる。

表4.6から表4.9より、動画全体と動画一部分における許容フレームに1人と2人がピークを持つときの回数がわかる。この回数が多いとき、周りの人物と違うタイミングで動作をしている場合と決められた動作をしていない場合が考えられる。

## 第5章 結論

### 5.1 まとめ

本研究では、複数人が同様のダンスを行う動画を対象として、各人物の動作のタイミングの一致度を評価した。まず、OpenPose から得られた各キーポイントの座標データからピークを求めた。動画内の人物の過半数が $K$ フレーム間にピークを持つときを正解の動作とする。このとき、各人物が正解の動作をした回数と正解の動作をしなかった回数を求め、各人物の動作の正解率を示した。これにより、各人物が決められた動作をしているかを評価した。また、各人物の正解率の大きさにより、評価対象者のダンス習熟度を評価した。次に、 $K$ フレーム間に動画内の過半数に満たない人物がピークを持つ回数を求めた。

### 5.2 課題

本研究では数値での評価はできたが、実際に動画を見た人による主観評価と数値の関係が不透明なため、動画の主観評価を含めた追加実験を行う必要がある。

また、各人物の手法 1 の結果により動画内の人物の習熟度を評価した。しかし、習熟度と正解率の関係を評価する方法を検討する必要がある。

## 謝辞

本研究の実験環境を与えてくださり、研究の方向性等の丁寧かつ熱心なご指導を頂いた渡辺教授に心から感謝いたします。

本研究のきっかけを与えてくださり、様々なご提案を頂きました早稲田大学国際情報通信センターの石川孝明様に心から感謝いたします。

日頃から御意見やアドバイスをくださった研究室の皆様に御礼申し上げます。

最後に、私をここまで育ててくださった家族に感謝いたします。

## 参考文献

- [1] 文部科学省, 新学習指導要領に基づく中学校向け「ダンス」リーフレット,  
[http://www.mext.go.jp/a\\_menu/sports/jyujitsu/1306098.htm](http://www.mext.go.jp/a_menu/sports/jyujitsu/1306098.htm), May, 2011
- [2] 南里卓也, 大津展之: “複数人動画像からの異常動作検出”,  
<http://css.risk.tsukuba.ac.jp/kashin/papers/2-1/PRMU04-Nanri.pdf>,
- [3] Z.Cao,T.Simon,S-E Wei,Y.Sheikh,“Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”,<https://arxiv.org/abs/1611.08050>,2016.
- [4] GitHub, inc. CMU-Perceptual-Computing-Lab/openpose,  
<https://github.com/CMU-Perceptual-Computing-Lab/openpose/blob/master/examples/media>, 24, April, 2017

## 図一覧

図 2.1	入力画像[4]	3
図 2.2	出力画像[4]	4
図 2.3	誤検出の例[4]	5
図 4.1	使用動画の第 1 フレーム	10
図 4.2	動画全体の 5 人の右手首 y 座標の変化	11
図 4.3	動画一部分の 5 人の右手首 y 座標の変化	11

## 表一覧

表 4.1	実験条件 .....	10
表 4.2	動画全体の各キーポイントの正解率 .....	12
表 4.3	動画全体の腕・脚部分と全身の正解率 .....	12
表 4.4	動画一部分の各キーポイントの正解率 .....	13
表 4.5	動画一部分の腕・脚部分と全身の正解数と不正解数 .....	13
表 4.6	動画全体の各キーポイントのピークの人数ごとの回数 .....	14
表 4.7	動画全体の腕・脚部分のピークの人数ごとの回数 .....	14
表 4.8	動画一部分の各キーポイントのピークの人数ごとの回数 .....	15
表 4.9	動画一部分の腕・脚部分のピークの人数ごとの回数 .....	15

## 研究業績

[1] 稲田, 石川, 渡辺, “OpenPose を用いた複数人のダンスの一致度評価”, 情報処理学会第 80 回全国大会, March, 13. 2018(発表予定)