

2025 年度（令和 6 年度）

創造工学セミナーⅡ Final Report

関節角度による運動姿勢の評価

研究メンバー

S521024 久留島 崇史

指導教員

金丸 隆志 教授

所属研究室

知能機械研究室

目次

第1章 緒論.....	4
1.1 研究背景.....	4
1.1.1 運動の需要.....	4
1.1.2 フィットネスクラブ会員数の増加.....	6
1.1.3 ボディビル.....	7
1.2 本研究の目的.....	8
第2章 先行研究.....	9
2.1 ダイビングスキルの定着を促す MediaPipe を使用した動作比較システム.....	9
2.2 関節角度による運動姿勢の評価システム.....	10
2.2.1 関節角度による運動姿勢の評価システムの目標.....	10
2.2.2 Web カメラを用いた関節角度による運動姿勢の評価システムの実験概要.....	12
2.2.3 Web カメラを用いた関節角度による運動姿勢の評価システムの実験結果.....	13
2.2.4 MediaPipe.....	14
2.2.5 2次元の関節角度算出手法.....	16
2.2.6 関節角度の計算手法.....	18
2.2.7 評価対象のポーズ.....	19
2.2.8 ポーズの判定に用いる角度.....	20
2.2.9 機械学習の概要.....	22
2.2.10 座標から角度への変換.....	23
2.2.11 良いポーズ・悪いポーズの定義.....	24
2.2.12 機械学習の条件.....	27
2.2.13 Web カメラを用いた関節角度による運動姿勢の評価システムの問題点.....	28
第3章 奥行きを考慮したポーズ評価システム.....	29
3.1 奥行きを考慮したシステムの構築.....	29
3.1.1 RealSense.....	29
3.1.2 スケールの統一.....	31
3.1.3 関節角度算出手法.....	32
3.1.4 Excel ファイルに記録するデータ.....	34
3.1.5 奥行きを考慮した良いポーズ・悪いポーズ.....	35
3.1.6 RealSense を用いても対応しないポーズ.....	38
3.1.7 3次元のポーズ評価システムにおける学習の条件.....	39
3.2 3次元のポーズ評価システムの精度検証実験.....	40
3.2.1 実験に用いた PC.....	40
3.2.2 実験方法.....	41

3.2.3 ポーズ評価システムの精度検証実験結果	43
第4章 結論	47
参考文献	48
謝辞	50

第1章 緒論

本章では、本研究の背景および目的について述べる。

1.1 研究背景

本節では研究に至った背景について述べる。

1.1.1 運動の需要

コロナ禍以降、リモートワークやオンライン授業が増え、外出機会が減っている。それにより運動不足を感じる人が多くなった。

スポーツ庁が行った「令和5年度「スポーツの実施状況等に関する世論調査」の概要」によると、令和5年度は20歳以上の日本人の52%が週に一回以上のスポーツをしていることがわかる[1]。

図1は概ね3年ごとに、平成24年度まで実施されてきた「体力・スポーツに関する世論調査」と平成27年度の「東京オリンピック・パラリンピックに関する世論調査」、平成29年度から1年ごとに行われている「スポーツの実施状況等に関する世論調査」を用いた20歳以上のスポーツ実施率の推移を示す図である。平成29年度からスポーツ実施率は50%台をキープしており、令和5年度においても、52%と半分以上の国民がスポーツを実施していることがわかる。

また、表1は令和五年度の調査で週に一回以上スポーツをすると回答した人に聞いた、スポーツを実施する理由を回答数上位から順に並べたものである。スポーツを実施する理由として「健康のため」、「体力増進、維持のため」、「運動不足を感じるから」が多く、現代の日本人は健康意識が高いことがわかる。

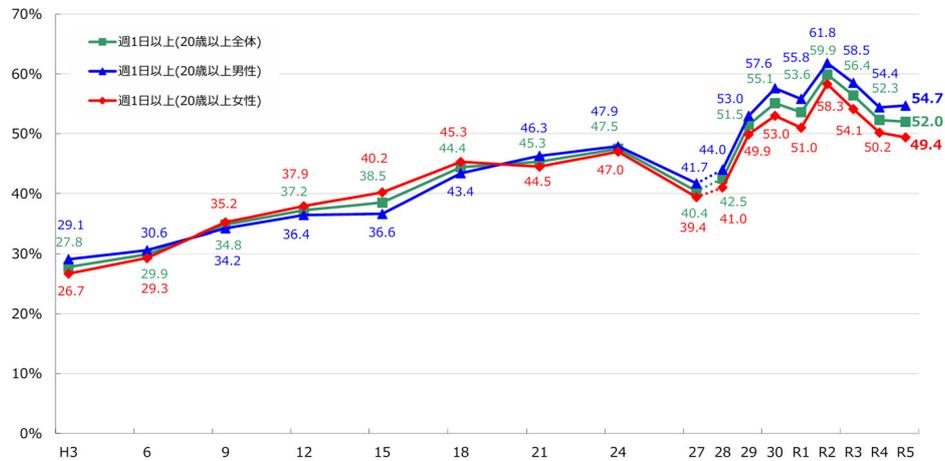


図 1 20歳以上のスポーツ実施率の推移 (H:平成、R:令和) [1]

表 1 運動・スポーツを週に1日以上実施した理由(令和5年度)[1]

運動・スポーツを週に1日以上実施した理由[%]	
健康のため	78.7
体力増進・維持のため	55.1
運動不足を感じるから	43.4
筋力増進・維持のため	38.6
楽しみ、気晴らしのため	38.3
肥満解消、ダイエットのため	28.9
有人・仲間との交流として	13.6
美容のため	9.7
自分の記録や能力を向上させるため	7.9
精神の修養や訓練のため	7.8
家族のふれあいとして	6.4

1.1.2 フィットネスクラブ会員数の増加

運動習慣のある人が増え、フィットネスクラブの会員に登録している人が増えている。図2は[2]で公開されているデータの「フィットネスクラブに個人で会員登録している人」の推移である。新型コロナウイルスの影響で2019年～2021年まで会員数が減ったが、それ以降の3年間で登録者数が増加していることがわかる。

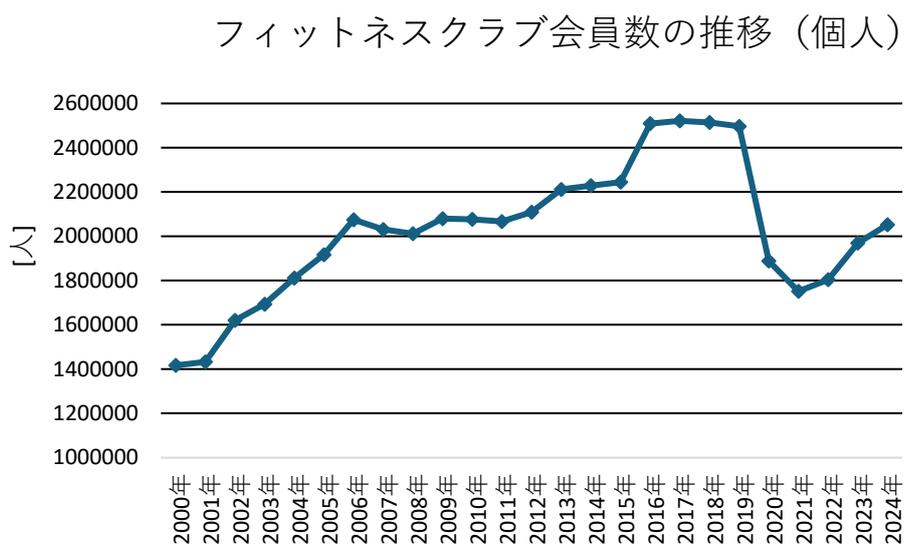


図2 フィットネスクラブ登録者数の推移[2]

1.1.3 ボディビル

フィットネスクラブなどでの運動の需要が増えると同時に、筋力トレーニング（以後筋トレと呼称）に注目が集まり、ボディビルにおいても同様のことがいえる。ボディビルでは、肉体を競い合うため、筋肉を強調する適切なポーズをとることが重要である。図 3 のポーズは腕の筋肉を強調できているが、図 4 のポーズは腕の筋肉を強調できていない[3]。

そのため、このポーズ（ダブルバイセップス）の姿勢を評価するシステムを実現することにした。

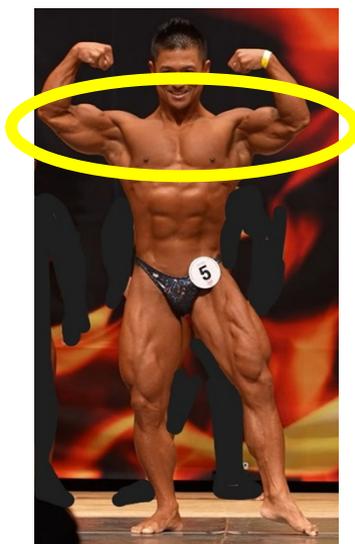


図 3 腕の筋肉を強調しているポーズ[3]



図 4 腕の筋肉を強調できていないポーズ[3]

1.2 本研究の目的

後述する先行研究で青井らはダブルバイセップスを評価するシステムを開発し、高い精度を示した[4]。また、その研究の中で奥行きを考慮して姿勢を評価できるシステムの研究も行っていた。しかし、その精度までは検証せず、実用性のあるシステムの実現には至らなかった。

そこで本研究の目的は、関節角度を用いたポーズ評価システムにおいて、奥行き情報を考慮した3次元ポーズ評価が、撮影条件の変化に対して有効であるかを検証することである。特に、ボディビルのダブルバイセップスを対象とし、2次元ポーズ評価システムと3次元ポーズ評価システムの判定結果を比較することで、奥行きを考慮したポーズ評価の有効性を明らかにすることを目的とする。

第2章 先行研究

本章では、本研究に関連する先行研究を紹介する。

2.1 ダイビングスキルの定着を促す MediaPipe を使用した動作比較システム

加藤らは、ダイビング技術の習得を支援することを目的として、MediaPipe を用いた動作比較システムを提案している。提案手法では、未熟練ダイバーとインストラクターの骨格情報を MediaPipe により取得し、両者の関節角度の差異を算出・可視化することで、姿勢の違いを定量的に評価している（図 5）[5]。

実験では、提案システムを用いたグループと従来の指導法のみを用いたグループを比較し、提案システムを使用したグループにおいて技能定着度が有意に向上することを確認している。また、提案システム使用前後も姿勢改善が認められ、この手法が技能習得支援に有効であることが示されている[5]。

一方で、この研究では撮影した映像データを一度 PC に転送して解析を行っており、リアルタイムでの動作比較は実現されていない。また、関節角度の差異は数値として表示されるものの、その値とスキルの良否との対応が直感的に理解しにくいという課題がある。

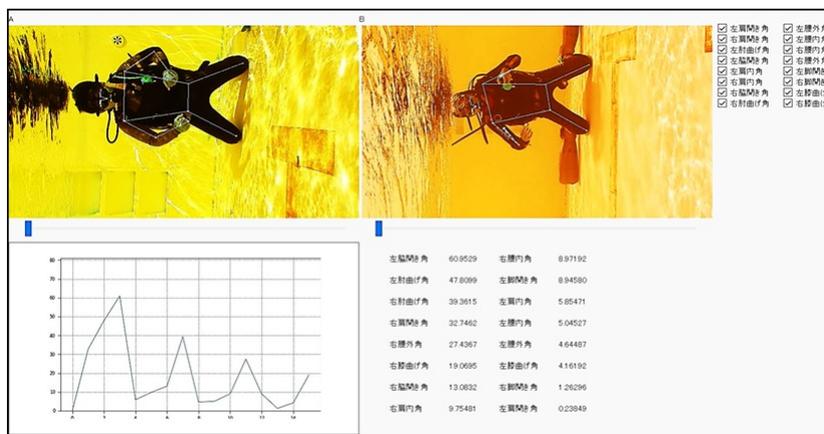


図 5 動作比較システムの表示画面構成[5]

2.2 関節角度による運動姿勢の評価システム

2.2 節では青井・池田の研究結果[4]をまとめる。

2.2.1 関節角度による運動姿勢の評価システムの目標

青井・池田は、前節の課題を踏まえ、ボディビルにおいてダブルバイセップスで効果的に筋肉を強調した姿勢が出来ているか評価する評価システムの構築を目標に研究を行った。このシステムは以下の3つの条件を満たすものとする。

1. 関節角度からボディビルのポーズの良し悪しを評価すること
2. リアルタイムで評価を得ること
3. 評価結果が直感的にわかること

図 6 は青井らが開発したボディビルのポーズのポーズ評価を行うシステムが正しく腕の筋肉を強調するポーズをとれていると評価した画像である。このような正しいと評価された姿勢に関しては、画面上で class0 と示される。

また図 7 は正しく腕の筋肉を強調するポーズをとれていないと評価された画像である。このような正しくないと評価された姿勢に関しては、画面上で class1 と示される。

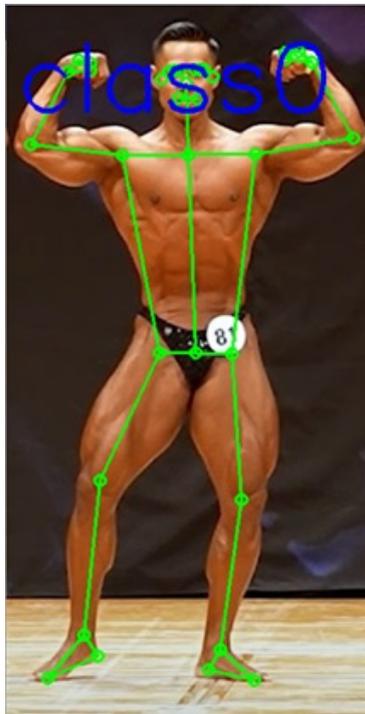


図 6 2次元で判定するシステムの判定結果（良いポーズ）[4]



図 7 2次元で判定するシステムの判定結果（悪いポーズ） [4]

2.2.2 Web カメラを用いた関節角度による運動姿勢の評価システムの実験概要

青井らが行った実験の方法は以下のとおりである。さらに、図 8 にその流れが図示されている。

I. 学習用画像の用意と学習

→良いポーズ(プロボディビルダー) 100 枚、悪いポーズ 200 枚のファイルを学習する。

II. 判定する画像の用意

→判定結果の評価に、学習データとは別の画像を 100 枚用意する。
(良いポーズと悪いポーズ各 50 枚)

III. 判定する画像の分類

→II で用意した画像を学習済みシステムに与え、良いポーズであれば class0、悪いポーズであれば class1 と判定結果を表示する。

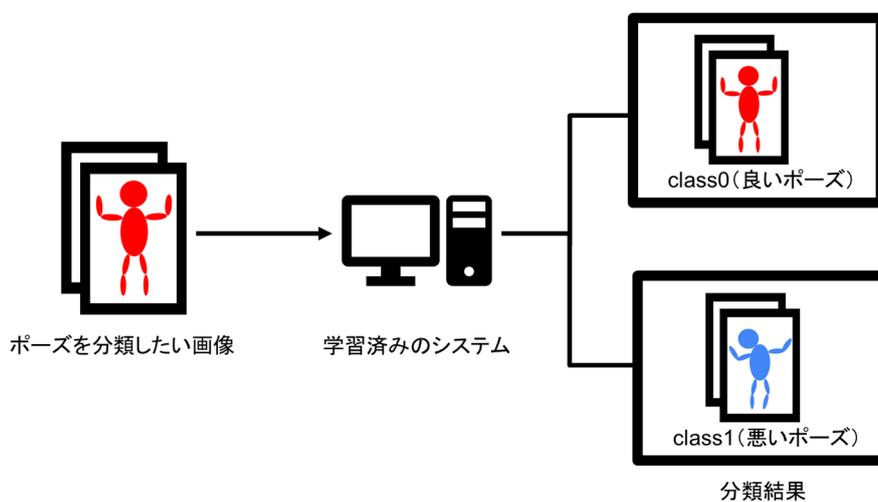


図 8 コンピューターによる運動姿勢の評価手法[4]

2.2.3 Web カメラを用いた関節角度による運動姿勢の評価システムの実験結果

2.2.2 節で行った実験の結果が以下である。全体の正答率が 90%となり、関節角度から姿勢を評価することが出来ているといえる結果が示された。

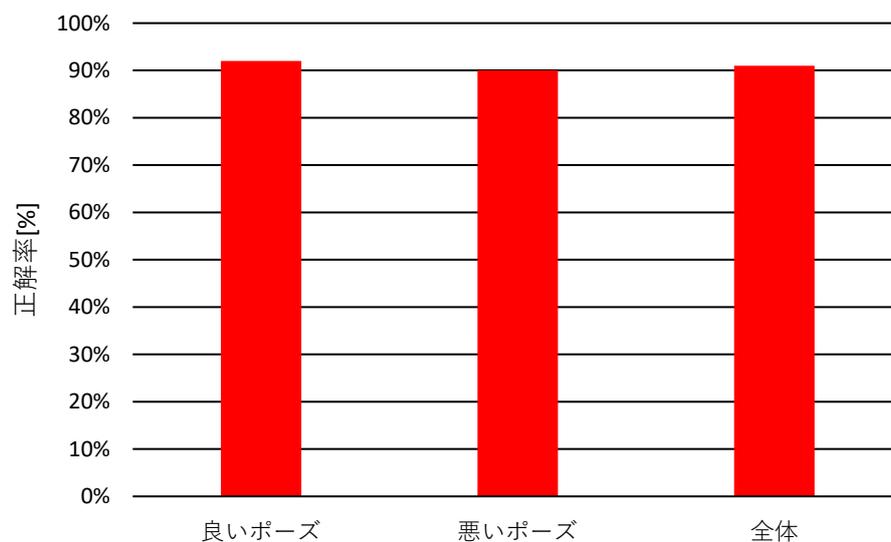


図 9 実験結果[4]

表 2 分類結果[4]

		予測	
		良いポーズ	悪いポーズ
結果	良いポーズ	46	5
	悪いポーズ	4	45

2.2.4 MediaPipe

関節角度による運動姿勢の評価システムに用いたライブラリである MediaPipe について説明する。MediaPipe とは Google が開発・提供しているリアルタイム向けの機械学習処理フレームワークである。主に画像・動画・ストリーミングデータを入力として、人の姿勢や手、顔などの特徴点を高速に推定することを目的として設計されている。

青井らは MediaPipe のうち、姿勢（ポーズ）を認識する機能を用いた[6]。ポーズ認識により、図 10 及び表 3 に示した 33 箇所の骨格点が推定される。このポーズ認識を利用するため、Kazuhito00氏が作成したサンプルプログラムを利用する[7]。認識結果が図 11 である。

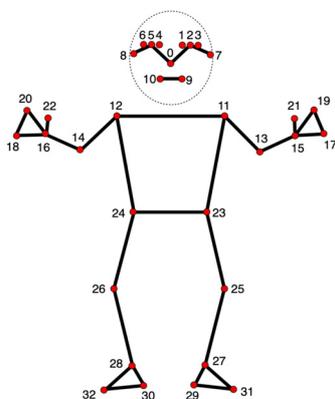


図 10 推定される骨格点の位置[6]

表 3 骨格点の名称[6]

番号	ランドマーク	番号	ランドマーク
0	nose	16	right wrist
1	left eye (inner)	17	left pinky
2	left eye	18	right pinky
3	left eye (outer)	19	left index
4	right eye (inner)	20	right index
5	right eye	21	left thumb
6	right eye (outer)	22	right thumb
7	left ear	23	left hip
8	right ear	24	right hip
9	mouth (left)	25	left knee
10	mouth (right)	26	right knee
11	left shoulder	27	left ankle
12	right shoulder	28	right ankle
13	left elbow	29	left heel
14	right elbow	30	right heel
15	left wrist	31	left foot index
		32	right foot index

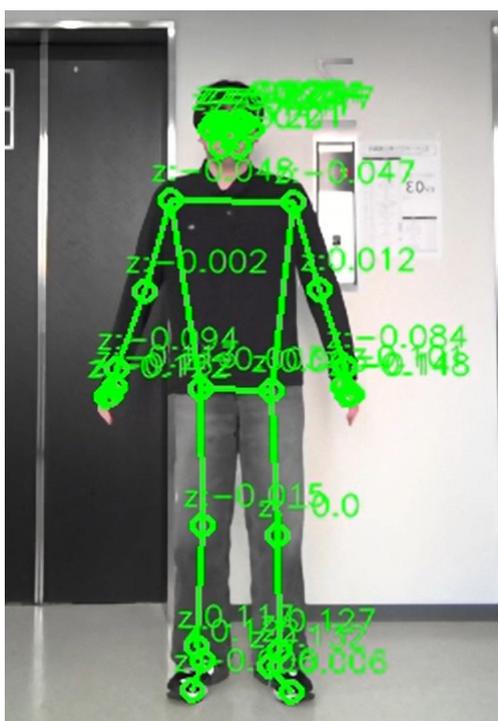


図 11 サンプルプログラムを用いて推定したポーズ

2.2.5 2次元の関節角度算出手法

MediaPipe で検出される骨格点の座標は二次元のピクセルの位置で取得できる。取得された 33 個の骨格点の位置は、図 12 のように画面左上を原点とした 2 次元座標として表現される。表 4 はそれをエクセルに出力したものである。(x,y) と同時に各骨格点の確からしさを示す信頼性も記録される。

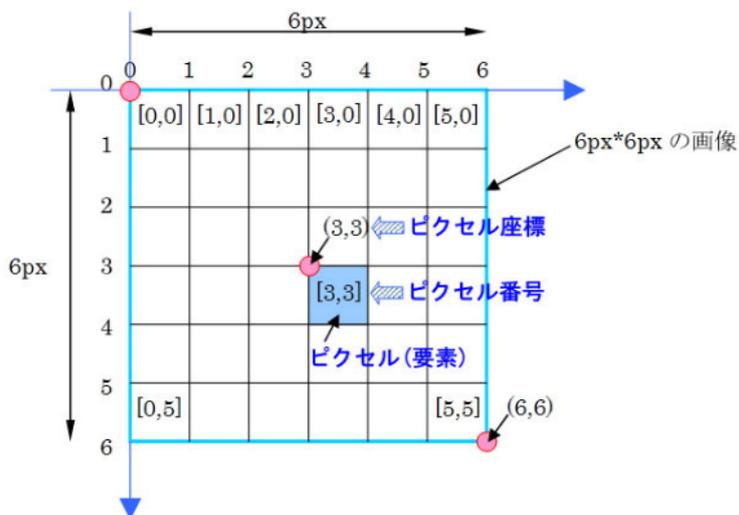


図 12 ピクセルで表される位置について[8]

表 4 検出された骨格点の座標と信頼性

	A	B	C
1	279	213	1
2	283	205	1
3	285	205	1
4	289	205	1
5	273	204	1
6	269	204	1
7	265	205	1
8	292	211	1
9	258	211	1
10	284	224	1
11	272	224	1
12	316	258	1
13	231	257	1
14	377	249	1
15	169	257	1
16	335	220	0.97
17	218	223	0.97
18	323	217	0.86
19	232	218	0.89
20	319	220	0.84
21	236	221	0.87
22	321	223	0.8
23	233	224	0.84
24	304	396	1
25	252	401	1
26	327	488	1
27	247	490	1
28	338	572	0.99
29	242	574	1
30	332	584	0.94
31	249	586	0.93
32	364	602	0.99
33	220	604	0.99
	X	Y	信頼性

2.2.6 関節角度の計算手法

青井らは、骨格点の座標から関節角度を求めるために、ベクトルの内積による角度算出手法を用いた。本節では、その計算方法について説明する。

ある骨格点 2 つを

$$\vec{a} = (x_1, y_1)$$

$$\vec{b} = (x_2, y_2)$$

とする。これら 2 つのベクトルのなす角を θ とすると、内積の定義より、次式が成り立つ。

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| |\vec{b}|}$$

ここで、内積およびベクトルの大きさを成分で表すと、上式は次のように書き換えられる。

$$\cos \theta = \frac{x_1 x_2 + y_1 y_2}{\sqrt{x_1^2 + y_1^2} \sqrt{x_2^2 + y_2^2}}$$

したがって、2 つのベクトルがなす角 θ は、逆余弦関数を用いて以下の式で求められる。

$$\theta = \cos^{-1} \left(\frac{x_1 x_2 + y_1 y_2}{\sqrt{x_1^2 + y_1^2} \sqrt{x_2^2 + y_2^2}} \right) [\text{rad}]$$

さらに角度の単位を度数法で扱うため、ラジアンから度への変換を行う。これにより、最終的な関節角度は次式で表される。

$$\theta = \frac{180}{\pi} \cos^{-1} \left(\frac{x_1 x_2 + y_1 y_2}{\sqrt{x_1^2 + y_1^2} \sqrt{x_2^2 + y_2^2}} \right) [^\circ]$$

以上の式を用いて、右肩・右肘・右手首の各骨格点の座標からベクトルを定義し、右肘の関節角度を算出した。それを、映像上に表示したのが図 13 である。このような角度を、姿勢評価に利用することができる。

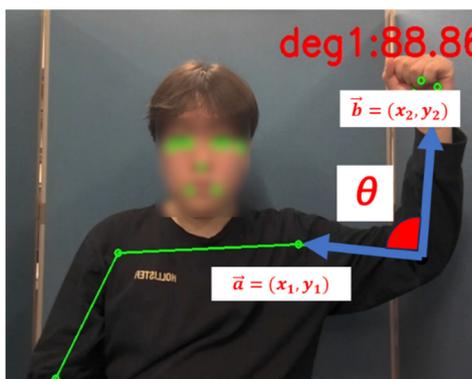


図 13 2次元ベクトルで算出された右肘の関節角度（鏡像）

2.2.7 評価対象のポーズ

本節では、青井らが評価対象としたボディビルのポーズについて説明する。

彼らは、図 14 のようなダブルバイセップスと呼ばれるポーズを対象とした。このポーズでは肘の角度、肘の位置、全身のバランス、大腿四頭筋の 4 点が評価される。それらを満たすことを条件としたポーズ評価を目標とした。図 15 は図 14 の写真で骨格を推定した例である。



図 14 ダブルバイセップスの例[3]

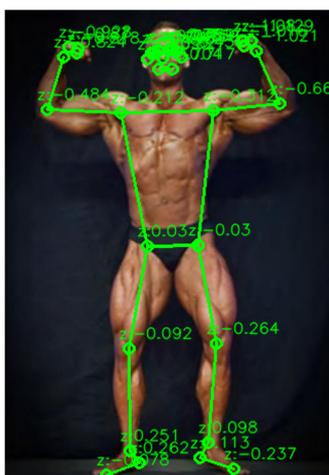


図 15 図 14 の骨格推定

2.2.8 ポーズの判定に用いる角度

本節では、ポーズの良否を判定するために用いる関節角度について説明する。

座標を判定に用いる場合、撮影画像内における被験者の立ち位置や体格の違いによって基準が変動し、同一のポーズであっても異なるデータとして扱われる可能性がある。その結果、ポーズの特徴を適切に判定できないことが想定される。そこで青井らは、関節角度を特徴量として採用した。

青井らの研究では、2.2.7項で示した評価ポイントを基準として、姿勢の特徴を表す角度を8種類選定した。選定した角度と対応する身体部位を表5と図16に示す。

これらの角度のうち、 θ_5 および θ_6 を算出するためには、身体全体の姿勢を表す基準線が必要となる。そのため、左右の肩および左右の腰のそれぞれの中点を算出し、それらを結んだ直線を身体を中心線として定義した。この中心線は、MediaPipeによって推定された骨格点をもとに新たに導入したものであり、身体の傾きや左右のバランスを評価するために用いている。

表5 ポーズの判定に用いる角度[4]

判定対象の箇所	角度
左肘	θ_1
右肘	θ_2
左脇	θ_3
右脇	θ_4
身体を中心線に対しての肩の傾き	θ_5
身体を中心線の傾き	θ_6
左膝	θ_7
右膝	θ_8

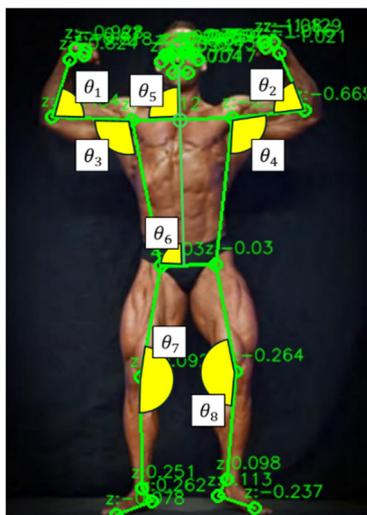


図 16 ポーズの判定に用いる角度

2.2.9 機械学習の概要

機械学習とは、ある目的を達成するための手順を予め用意されたデータに基づき自動的に決定する技術である[9]。機械学習を用いた理由は関節角度を基準にボディビルのポーズの良し悪しを評価するための学术论文などのデータが数少ないためである。そのためプロがとるポーズから関節角度を抽出し、コンピューターに学習させることで問題の解決を図った。

機械学習の流れは図 17 のとおりである。

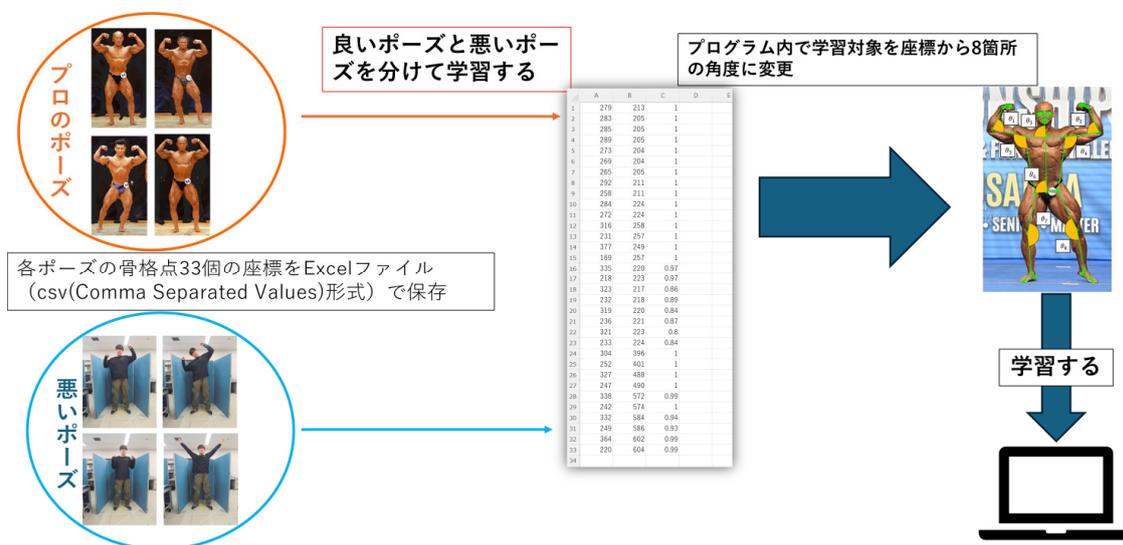


図 17 機械学習の流れ

2.2.10 座標から角度への変換

2.2.8 で述べたように、青井らの研究では、姿勢の特徴量として骨格点の座標そのものではなく、関節角度を用いた。入力データは、MediaPipe により推定された 33 箇所の骨格点の 2 次元座標から算出された 8 種類の関節角度なのだった。

骨格データは、33 箇所の骨格点それぞれに対応する 2 次元座標として取得される。このとき、入力ベクトルは次のように表される。

$$\mathbf{X} = [[x_1, y_1], [x_2, y_2], \dots, [x_{33}, y_{33}]]$$

次に、これらの座標情報から、2.2.8 節で定義した 8 種類の関節角度 $\theta_1 \sim \theta_8$ を算出し、姿勢を表す特徴量ベクトルへと変換する。

$$\mathbf{X} = [\theta_1, \theta_2, \dots, \theta_8]$$

この処理により、1 つの姿勢データは 8 次元の角度特徴量として表現される。

さらに、学習データを i 個用意した場合、入力データ全体は以下のような行列で構成される。

$$\mathbf{X} = \begin{bmatrix} [\theta_{1(1)}, \theta_{2(1)}, \dots, \theta_{8(1)}] \\ [\theta_{1(2)}, \theta_{2(2)}, \dots, \theta_{8(2)}] \\ \vdots \\ [\theta_{1(i)}, \theta_{2(i)}, \dots, \theta_{8(i)}] \end{bmatrix}$$

各姿勢データに対応する正解ラベル（ターゲット）は、以下のようなベクトルとして与えられる。

$$\mathbf{y} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(i)} \end{bmatrix}$$

以上のように、本研究では座標情報を関節角度に変換することで、姿勢の特徴を抽出し、機械学習に適した入力データを構成している。

2.2.11 良いポーズ・悪いポーズの定義

青井らの研究ではプロが実際の大会で作ったポーズを良いポーズ(class0)、青井らが作った2.2.7節の4つの評価ポイントを満たしていないポーズを悪いポーズ(class1)と定義した。

良いポーズの具体例が図18であり、悪いポーズの具体例が図19、図20、図21、図22である。また表6は悪いポーズの画像がそれぞれどの評価ポイントを満たしていないかを示している。

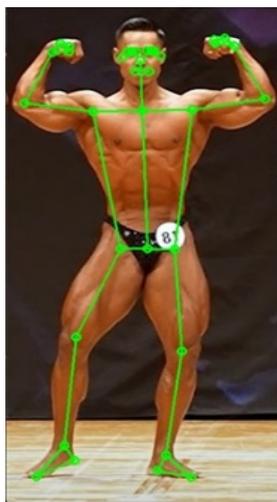


図18 良いポーズ(class0)[4]



図19 悪いポーズ1(class1)[4]



図 20 悪いポーズ 2(class1)[4]



図 21 悪いポーズ 3(class1)[4]



図 22 悪いポーズ 4(class1)[4]

表 6 ポーズと評価ポイントの対応[4]

	肘の角度	肘の位置	全身のバランス	大腿四頭筋
図 18	○	○	○	○
図 19	×	○	○	○
図 20	×	×	○	○
図 21	○	×	×	○
図 22	○	○	○	×

2.2.12 機械学習の条件

青井らは、良いポーズ 100 枚、悪いポーズ 200 枚の合計 300 枚の静止画像を用いて機械学習を行った。学習に用いた条件を表 7 に示す。

機械学習では、学習を何回繰り返すかを表す数値としてエポック数 (epoch) を設定する。

1 エポックとは、用意したすべての学習データを一度使って学習を行うことを意味する。青井らは、学習が十分に進むよう、エポック数を 1000 に設定した。

また、1 回の学習でまとめて使用するデータ数をバッチサイズ (batch size) と呼ぶ。青井らはバッチサイズを 50 とした。この設定により、300 枚の学習データは 1 エポックあたり 6 回に分けて学習に使用される。

このとき、1 回分の学習処理をイテレーション (iteration) と呼び、1 エポックあたりのイテレーション数は 6 回となる。各イテレーションでは、300 枚の画像の中から 50 枚がランダムに選ばれ、学習が行われる。

以上の条件より、1000 エポック \times 6 イテレーションとなり、合計で 6000 回の学習処理が行われることになる。

表 7 学習の条件[4]

エポック数	1000
バッチサイズ	50
イテレーション	6

2.2.13 Web カメラを用いた関節角度による運動姿勢の評価システムの問題点

青井らのシステムには問題点があった。それはカメラに対して被験者が正面に立っていない場合、正しい関節角度を算出することができず、正しく姿勢の判定ができない点である。

これは、青井らが使用していたカメラが Web カメラであり、取得できる情報が 2 次元に限られていることが原因である。そのため、奥行き方向の情報を含む姿勢では、骨格点の位置を正しく取得できず、関節角度の計算に誤差が生じるという問題点がある。

奥行きを含む姿勢を正しく評価するためには、3次元の情報を取得できるカメラを用いる必要があり、本研究では RealSense を用いることにする。

第3章 奥行きを考慮したポーズ評価システム

本章では、青井・池田らが残したプログラム[4]をもとに本研究で構築したシステムとその実験結果についてまとめる。

3.1 奥行きを考慮したシステムの構築

本節では、本研究で構築した、奥行き情報を考慮したポーズ評価システムの概要を述べる。

3.1.1 RealSense

本節では、本研究で使用する RealSense について説明する。

RealSense は、Intel 社が開発した深度カメラであり、一般的な Web カメラでは取得できない奥行き情報を取得できる点が特徴である。この奥行き情報を利用することで、人物の位置関係を 3 次元的に把握することが可能となる。

本研究では、RealSense D435 (図 23) を使用し、3 次元情報を用いたポーズ認識を行う。使用した RealSense の主な仕様を表 8 に示す。

RealSense では、通常の RGB 画像に加えて、深度センサから得られる Depth 画像を同時に取得することができる。RGB 画像の例を図 24 に、Depth 画像の例を図 25 に示す。本研究では、RealSense を USB 接続により PC と接続して使用した。

また、ポーズ認識は、第 2 章で 2 次元ポーズ評価に用いた MediaPipe を RGB 画像に対して適用することで実現する。取得した骨格情報に Depth 情報を加えてポーズ評価を行う。



図 23 RealSense D435

表 8 RealSense D435 の仕様

製品名	Intel RealSenseDepth Camera D435
デプス解像度	1280×720
RGB解像度	1920×1080
最大デプスFPS	90
デプス範囲	0.2[m]~10[m]
デプス対角画角	90[°]
寸法	90[mm]×25[mm]×25[mm]



図 24 RealSense で取得した RGB 画像

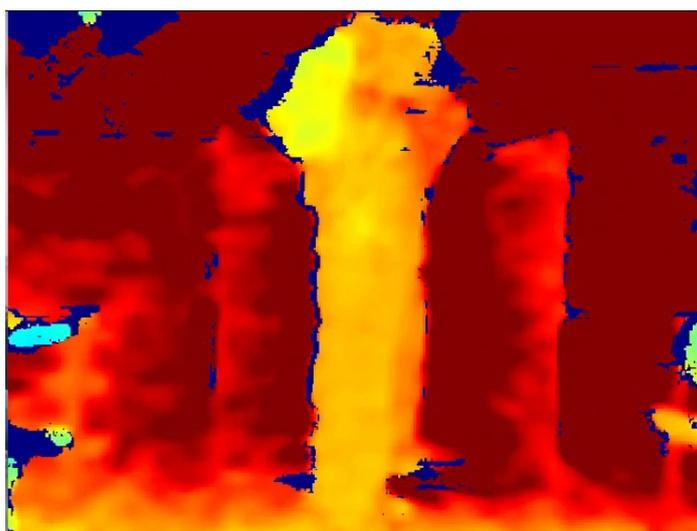


図 25 RealSense で取得した Depth 画像

3.1.2 スケールの統一

RealSense を用いた 3 次元での関節角度計算では、座標の扱い方に注意する必要がある [4]。RGB 画像上で取得される骨格点の位置はピクセルで表される一方、深度情報はカメラからの距離としてメートル単位で得られるためである (図 26)。

これらの情報をそのまま用いると方向ごとに単位が異なり、角度計算の結果が不安定になるため、関節角度を計算する前に、骨格点の位置をすべて同じ単位に変換する必要がある。

具体的には、画面上の位置 (x, y) と深度情報 z を組み合わせ、実際の空間上での位置を求める方法を用い、ピクセル座標と深度情報を 3 次元座標へ変換する [10]。

この変換を、RealSense の Python 用ライブラリ「pyrealsense2」に含まれる「rs2_deproject_pixel_to_point()」で実現した。この処理により、骨格点の位置は図 26 のように、すべてメートル単位で表される。

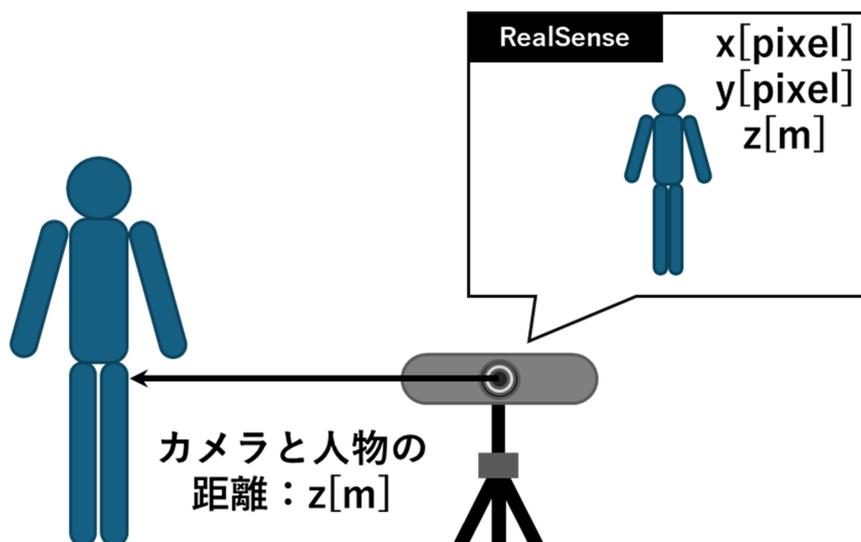


図 26 z 軸の定義

3.1.3 関節角度算出手法

本節では、3次元座標を用いて関節角度を計算する方法について説明する。

ここで用いる座標は、3.1.2節で述べた方法により、すべてメートル単位に変換されたものとする。

3次元空間における関節角度は、2次元の場合と同様に、関節をはさむ2つのベクトルのなす角として求める。本研究では、関節位置を基準として隣り合う骨格点の座標差からベクトルを作成し、その角度を計算する方法を用いた。

3次元空間上の2つのベクトルを

$$\vec{a} = (x_1, y_1, z_1)$$

$$\vec{b} = (x_2, y_2, z_2)$$

とする。このとき、2つのベクトルのなす角 θ は、内積を用いて次のように表される。

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| |\vec{b}|}$$

内積および各ベクトルの大きさを成分で表すと、上式は以下のように書き換えられる。

$$\cos \theta = \frac{x_1 x_2 + y_1 y_2 + z_1 z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2} \sqrt{x_2^2 + y_2^2 + z_2^2}}$$

したがって、角度 θ は逆余弦関数を用いて求められる。

$$\theta = \cos^{-1} \left(\frac{x_1 x_2 + y_1 y_2 + z_1 z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2} \sqrt{x_2^2 + y_2^2 + z_2^2}} \right) [\text{rad}]$$

本研究では、角度の表示を度数法で行うため、ラジアンから度への変換を行い、最終的に次式を用いて角度を算出した。

$$\theta = \frac{180}{\pi} \cos^{-1} \left(\frac{x_1 x_2 + y_1 y_2 + z_1 z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2} \sqrt{x_2^2 + y_2^2 + z_2^2}} \right) [^\circ]$$

この方法で右肘の角度を算出した様子を示したのが図 27、図 28 である。

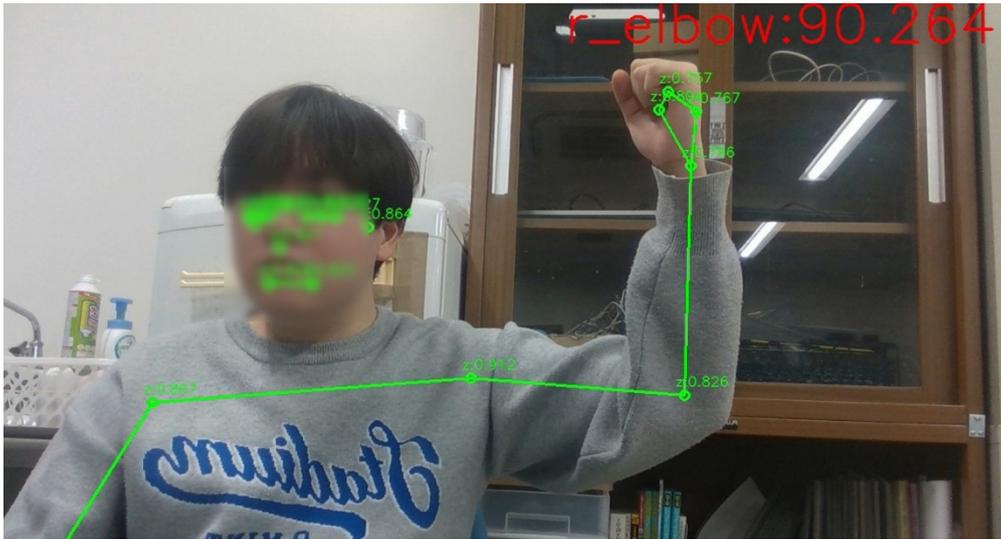


図 27 3次元座標を用いて導出した右肘の角度1 (鏡像)

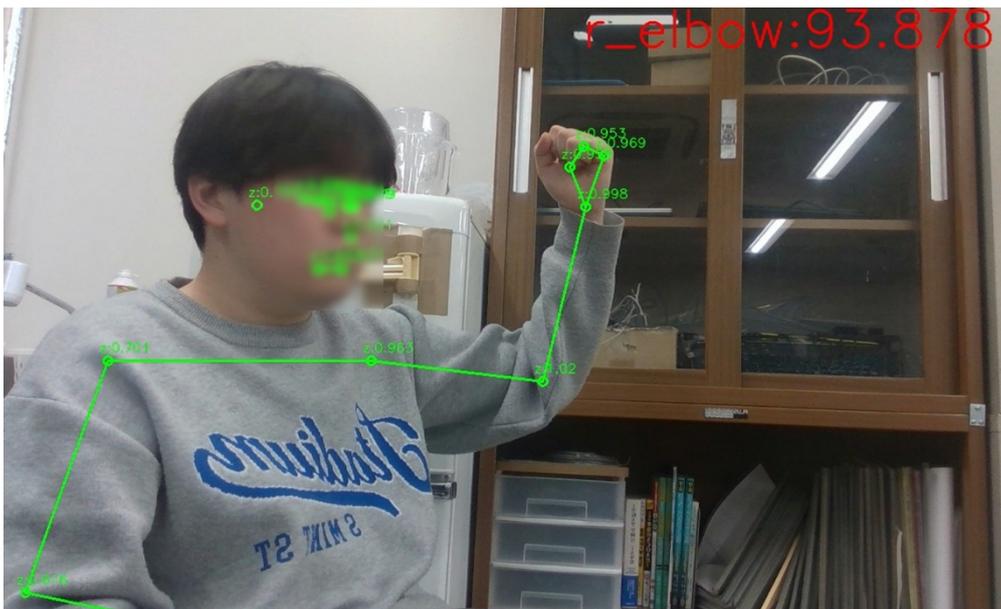


図 28 3次元座標を用いて導出した右肘の角度2 (鏡像)

3.1.4 Excel ファイルに記録するデータ

Web カメラを用いた2次元のポーズ評価システムの場合 Excel ファイル (csv 形式) に記録されるのは (x,y) のピクセル座標と信頼性からなる 33 点 \times 3=99 個のデータだったが、3次元の場合追加で (x,y) のメートル座標と $z[m]$ が加わり、データ数が $33\times 6=198$ 個となる。それを記録した Excel ファイルで学習を行う。

表 9 3次元の座標を記録する Excel の例

	A	B	C	D	E	F
1	556	126	2.44	-0.21	-0.64	1
2	560	115	2.45	-0.2	-0.67	1
3	564	115	2.45	-0.19	-0.68	1
4	567	115	2.44	-0.18	-0.67	1
5	548	115	2.41	-0.23	-0.66	1
6	543	116	2.41	-0.24	-0.66	1
7	538	116	2.39	-0.25	-0.66	1
8	572	121	2.45	-0.17	-0.66	1
9	525	124	2.41	-0.29	-0.64	1
10	563	140	2.44	-0.19	-0.6	1
11	547	140	2.53	-0.24	-0.63	1
12	612	182	2.77	-0.07	-0.56	1
13	489	188	2.6	-0.41	-0.51	1
14	691	160	2.36	0.14	-0.54	1
15	411	170	2.79	-0.68	-0.6	1
16	639	95	2.42	0.01	-0.72	0.99
17	467	104	2.4	-0.44	-0.69	0.99
18	627	83	2.11	-0.02	-0.65	0.95
19	479	90	2.35	-0.4	-0.71	0.94
20	619	86	2.51	-0.04	-0.77	0.94
21	487	94	2.31	-0.37	-0.69	0.93
22	619	93	2.45	-0.04	-0.73	0.92
23	486	101	2.41	-0.39	-0.7	0.91
24	597	369	2.32	-0.1	0	1
25	521	370	2.39	-0.29	0	1
26	628	513	2.52	-0.02	0.39	0.99
27	511	521	2.53	-0.34	0.42	0.99
28	643	645	2.51	0.02	0.75	0.98
29	498	656	2.5	-0.37	0.78	0.99
30	625	666	2.46	-0.03	0.79	0.92
31	510	674	2.51	-0.34	0.83	0.88
32	684	684	2.57	0.14	0.87	0.97
33	483	699	2.89	-0.48	1.03	0.98
...	x[px]	y[px]	z[m]	x[m]	y[m]	信頼性

3.1.5 奥行きを考慮した良いポーズ・悪いポーズ

2次元のポーズ評価システムにおいては良いポーズはプロのポーズを用いて学習を行った。

しかし、3次元の座標を取得するには、RealSense を用いる必要があるため、プロのポーズを学習させることはできない。

そこで、本研究では、筆者自身が 2.2.7 節で示した4つの評価ポイントを満たすポーズを RealSense で撮影し、良いポーズとして学習させる。また、悪いポーズも同様とする。

図 29、図 30 のような評価ポイントを満たすポーズを良いポーズ、図 31、図 32 のような評価ポイントを満たさないポーズを悪いポーズとした。

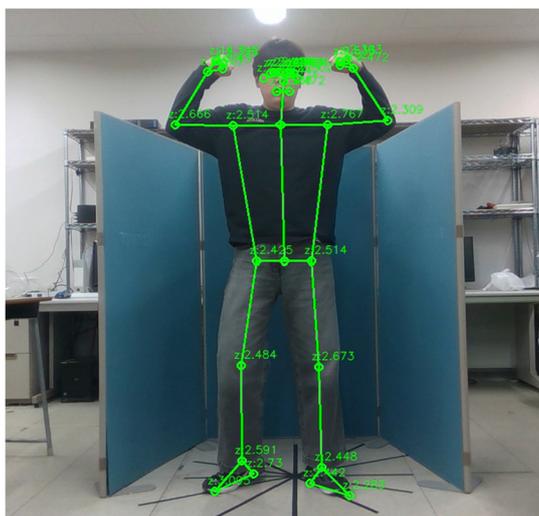


図 29 RealSense で撮影した良いポーズ 1 (class0)

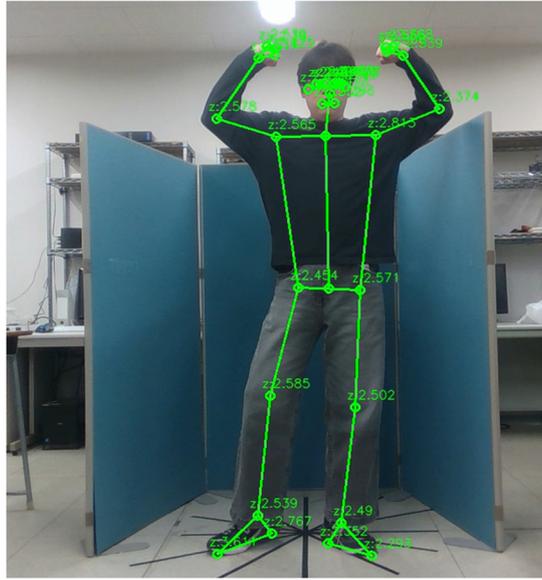


図 30 RealSense で撮影した良いポーズ 2 (class0)

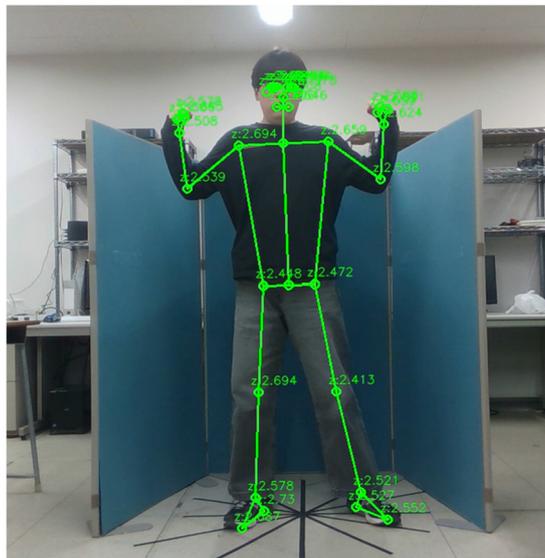


図 31 RealSense で撮影した悪いポーズ 1 (class1)

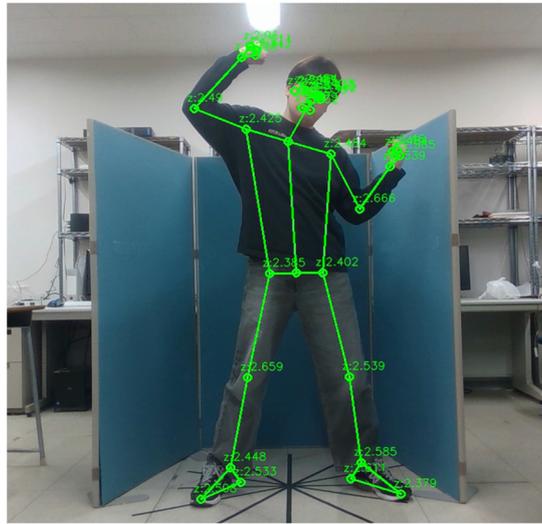


図 32 RealSense で撮影した悪いポーズ 2(class1)

3.1.6 RealSense を用いても対応しないポーズ

RealSense を用いることで奥行きを考慮したポーズの評価を行うことが出来る。しかし図 33、図 34 のようにカメラに対して被験者が $-90,90[^\circ]$ の向きをとっている時は、カメラの画角に写っていない骨格は認識できない。そのため、このような姿勢に対してはポーズ評価を行わないものとする。



図 33 被験者が平面角 $90[^\circ]$]

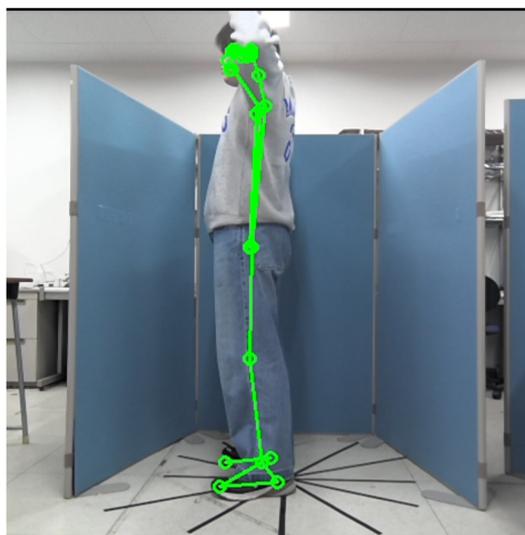


図 34 被験者が平面角 $-90[^\circ]$]

3.1.7 3次元のポーズ評価システムにおける学習の条件

3次元のポーズ評価の学習に用いる条件は以下のとおりである。

表 10 学習条件

エポック数	1000
バッチサイズ	100
イテレーション	5

3.2 3次元のポーズ評価システムの精度検証実験

本節では本研究で構築した「3次元のポーズ評価システム」の精度を実験した環境、方法、結果についてまとめる。

3.2.1 実験に用いたPC

実験に用いたPCの実行環境についてまとめたものが表 11 に示す。

表 11 本実験における使用環境

OS	Microsoft Windows 11 Home
CPU	Intel(R) Core(TM) i7-1065G7 CPU @ 1.30GHz (1.50 GHz)
RAM	32.0 GB

3.2.2 実験方法

本節では3次元のポーズ評価システムの実験方法について述べる。図 35 は実験環境の写真である。

I. ポーズの学習

RealSense を用い撮影した良いポーズ(筆者が 2.2.7 節の4つの評価ポイントを満たすよう撮影したもの) 100 枚(例:図 29、図 30)、悪いポーズ 200 枚(例:図 31、図 32) のファイルを学習する。

II. ポーズの判定

被験者が良いポーズ(実験者が 2.2.7 節の4つの評価ポイントを満たすよう撮影したもの) をしている姿をリアルタイムでシステムが判定する。良いポーズであれば class0、悪いポーズであれば class1 と判定結果が表示される。

奥行きを考慮したポーズを本システムが判定できるかを確認するため、カメラと被験者の角度を変えながら (0° , $\pm 30^\circ$, $\pm 60^\circ$) 判定を実施する(図 36)。

III. 判定の記録

判定結果の記録は、リアルタイムのポーズ評価システムで被験者が 2.2.7 節の4つの評価ポイントを満たしているポーズをとっている姿と、評価ポイントを満たしていないポーズをとっている姿を撮影し、姿勢の良し悪しをどちらに分類しているかを1秒ごとに合計 50 回の評価を目視で確認する。

上記の流れで精度を実験した。

また、青井らによる2次元のポーズ評価システムでも同様の学習データを用いた比較実験を行った。その場合もカメラは Web カメラではなく RealSense を用いた。ただし判定に使用した画像は RGB 画像のみであり、学習させたデータも (x, y) のピクセル座標のみである。



RealSense

図 35 実験環境の写真

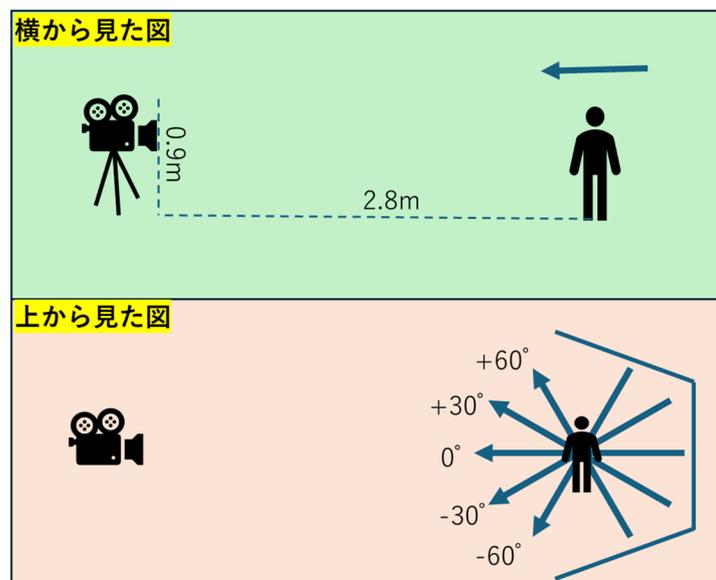


図 36 カメラと被験者の関係

3.2.3 ポーズ評価システムの精度検証実験結果

ポーズ評価システムの精度を検証する実験を行った。

図 37 および図 38 に、3次元ポーズ評価システムにおける正しい判定の割合（良いポーズ（class0）および良くない姿勢（class1））を示す。

また、図 39 および図 40 に、2次元ポーズ評価システムにおける正しい判定の割合を示す。

この結果から3次元ポーズ評価システムは撮影角度による影響を受けにくく、一定の判定結果を示す傾向があることが確認された。一方で、2次元ポーズ評価システムでは、カメラに対する被験者の角度に大きく依存する結果となった。

また表 12、表 13 はそれぞれ3次元ポーズ評価システムと2次元のポーズ評価システムの全角度の分類結果をまとめた行列である。

正しい判定の割合 (3D:class0)

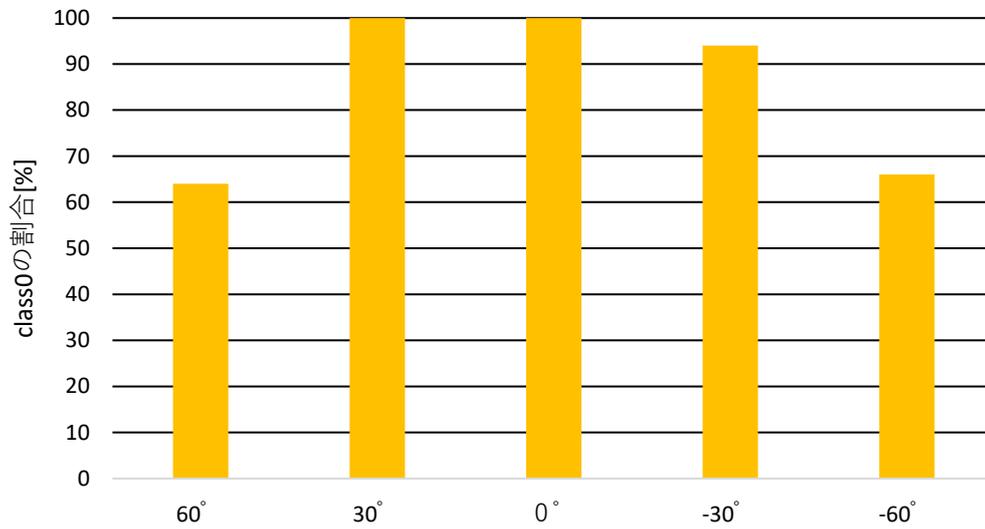


図 37 良いポーズに対しての3次元ポーズ評価システムの正答率

正しい判定の割合 (3D:class1)

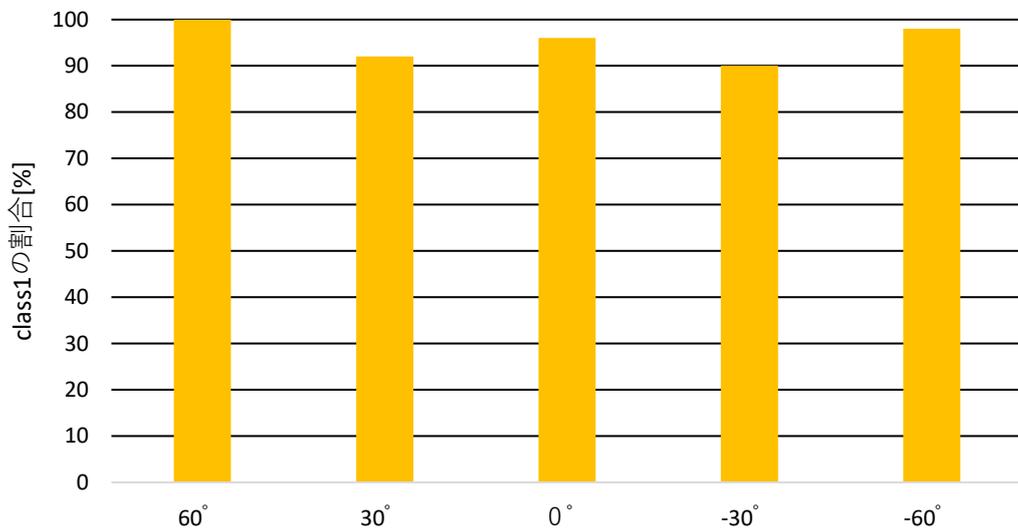


図 38 悪いポーズに対しての3次元ポーズ評価システムの正答率

正しい判定の割合 (2D:class0)

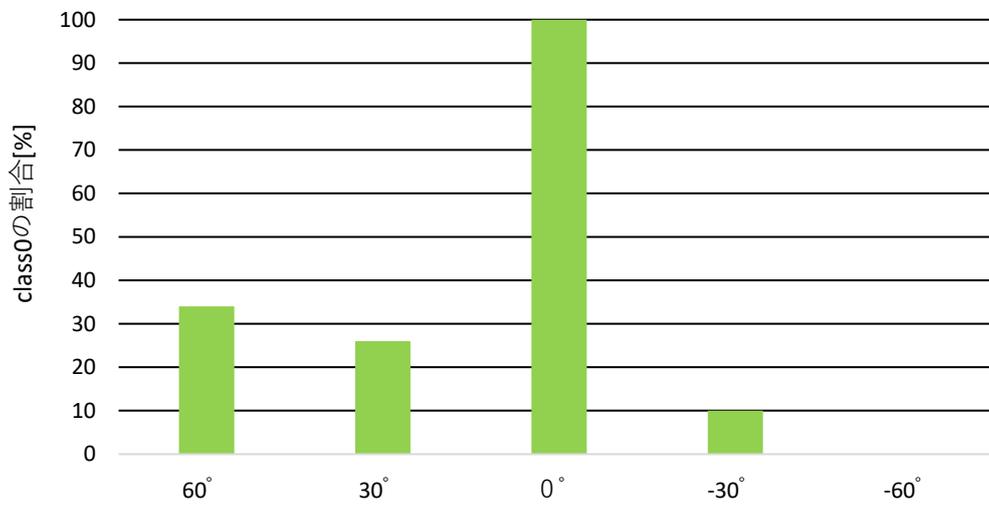


図 39 良いポーズに対しての2次元ポーズ評価システムの正答率

正しい判定の割合 (2D:class1)

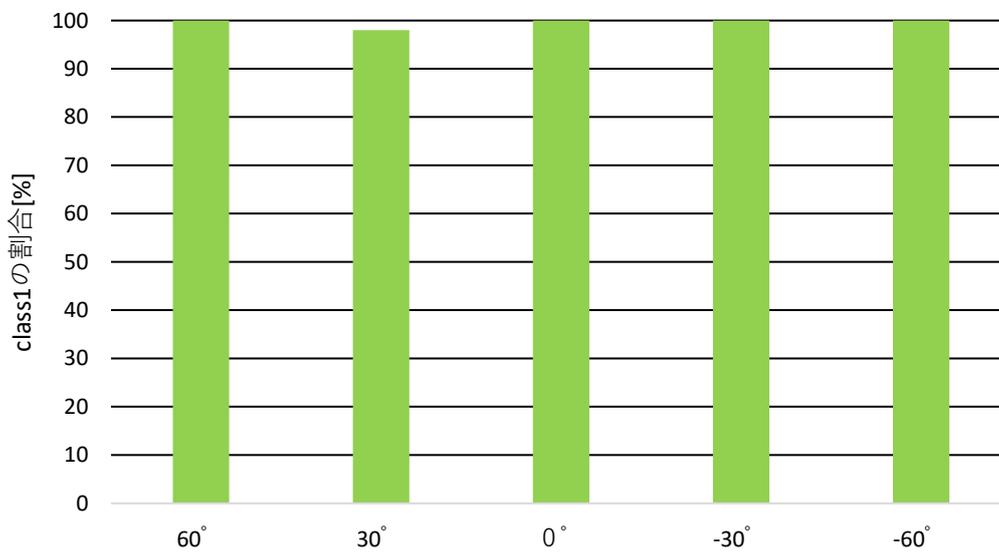


図 40 悪いポーズに対しての2次元のポーズ評価システムの正答率

表 12 3次元ポーズ評価システムの分類結果

3次元のポーズ評価システム		判定	
全角度合計		良いポーズ	悪いポーズ
実際	良いポーズ (class0)	212	38
	悪いポーズ (class1)	12	238

表 13 2次元ポーズ評価システムの分類結果

2次元のポーズ評価システム		判定	
全角度合計		良いポーズ	悪いポーズ
実際	良いポーズ (class0)	85	165
	悪いポーズ (class1)	1	249

第4章 結論

本研究では、関節角度を用いたポーズ評価システムにおいて、2次元ポーズ評価システムと3次元ポーズ評価システムの精度を比較し、奥行きを考慮したポーズ評価の有効性について検証を行った。

その結果、2次元ポーズ評価システムは、悪いポーズの判定において高い正答率を示す一方で、撮影角度が変化した場合に良いポーズを誤判定する傾向が強いことが確認された。また、3次元ポーズ評価システムでは、撮影角度が変化しても判定結果が大きく偏ることはなく、良いポーズおよび悪いポーズの両方を一定の基準で評価できる傾向が確認された。

以上の結果から、3次元ポーズ評価を用いることで、奥行きを考慮したポーズ評価が可能となり、撮影条件の変化に対して有効な評価手法である可能性が示された。

参考文献

- [1] 令和5年度「スポーツの実施状況等に関する世論調査」の概要 | スポーツ庁
https://www.mext.go.jp/sports/content/20240327-kensport01-000034684_1_2.pdf
(閲覧日 2026年2月2日)
- [2] 年・実数 Sheet (B9~B33,N9~N33) | フィットネスクラブ | 長期データ | 特定サービス産業動態統計調査 | 経済産業省
https://www.meti.go.jp/statistics/tyo/tokusabido/result/result_1.html
(閲覧日 2026年2月7日)
- [3] 審査員としての心構え | J B B F 競技ルールマニュアル | (公社)日本ボディビル・フィットネス連盟
https://www.jbbf.jp/download/Committee/2025%94N%93x%94%C5_JBBF%8B%A3%8BZ%83%8B%81%5B%83%8B%83%7D%83j%83%85%83A%83%8B.pdf
(閲覧日 2026年2月2日)
- [4] 青井 一馬、池田 周磨 “関節角度による運動姿勢の評価に関する研究” 工学院大学 2024年度卒業論文 (2025) .
(閲覧日 2026年2月2日)
“テスト”
- [5] 加藤 司、安富祖 仁 “ダイビングスキル定着を促す MediaPipe を使用した動作比較システムの開発” 電気学会論文誌 D,143 巻,2号,p.118-124,(2023).
https://www.jstage.jst.go.jp/article/ieejias/143/2/143_118/_article/-char/ja/
(閲覧日 2026年2月2日)
- [6] Google “姿勢ランドマーク検出ガイド” Google AI for Developers (2024/10/13).
https://ai.google.dev/edge/mediapipe/solutions/vision/pose_landmarker?hl=ja
(閲覧日 2026年2月2日)
- [7] Kazuhito00 “mediapipe-python-sample v0.8.9” GitHub (20-24/10/13).
<https://github.com/Kazuhito00/mediapipe-python-sample/releases/tag/v0.8.9>
(閲覧日 2026年2月2日)

[8] ガンマソフト株式会社 “ピクセル座標とは：グーグルマップのしくみを探る(2)”
maple.net (2011/01/11).

<https://www.mapli.net/blogs/google-maps-mechanism/how-google-map-works2/>

(閲覧日 2026 年 2 月 2 日)

[9] 金丸 隆志 “Raspberry Pi ではじめる機械学習～基礎からディープラーニングまで～”
株式会社講談社 (2019).

[10] keoitate:株式会社日立製作所 “pyrealsense2 で距離カメラのデータを扱う” Qiita
(2020/12/18).

<https://qiita.com/keoitate/items/efe4212b0074e10378ec>

(閲覧日 2026 年 2 月 7 日)

謝辞

本研究を進めるにあたり、2年間にわたって多大なるご指導とご助言を賜りました金丸隆志教授に、心より感謝申し上げます。研究の方向性や進め方について数多くのご指摘を頂いたにもかかわらず、私自身の理解不足や取り組みの遅れにより、最後まで十分に応えられない点も多く、度々お手数をおかけしました。改めて深く感謝申し上げます。ありがとうございました。