

カメラ画像を用いた体幹トレーニングの姿勢支援手法の提案

綿谷 惇史^{†1,a)} 謝 浩然^{†1,b)} 宮田 一乗^{†1,c)}

概要：近年、深層学習を用いて、ひとつの画像から姿勢を推定する研究が多く行われている。一方、体幹は運動やバランスに対して重要な役割を果たしており、体幹を鍛える体幹トレーニングに注目が集まっている。体幹トレーニングは一般的に個人でも行うことができるトレーニングであり、その効果を最大限に発揮するには、正しい姿勢の保持が極めて重要である。しかしながら、個人がトレーニング中の姿勢を把握することや正しい姿勢と判断するのは困難である。これまでに様々なトレーニングを対象とした姿勢支援は行われているが、深度カメラや複数台のカメラを必要とするため、コストがかかる。また、ユーザへの視覚的フィードバックが1視点または骨格情報のみを示している2次元ボーン画像であるため、姿勢の把握が難しい。そこで、本研究では、深層学習を利用して姿勢を推定する研究を応用し、単一RGBカメラのみを用いた体幹トレーニングの姿勢支援手法を提案する。提案手法では、カメラ画像から姿勢推定を行い、3次元モデルを生成し、目標姿勢と現在の姿勢を重畳表示する。さらに、目標姿勢と現在の姿勢の違いを把握しやすくするために、身体各部位10箇所（両手および肘、肩、膝、足首）にマーカを表示し、誤差に応じて、4段階に色変化させる。これらの情報を2視点でユーザへ視覚的フィードバックすることで、姿勢補正の支援を行う。

Trunk Training Support System from Camera Image

ATSUSHI WATATANI^{†1,a)} HAORAN XIE^{†1,b)} KAZUNORI MIYATA^{†1,c)}

Abstract: Pose estimation from a single image has been intensively explored by deep learning approaches recently. However, the applications based on these approaches are still in progress. In this work, we aim to propose a support user interface for trunk training from a single camera image. The trunk plays an important role in exercise and balance, and trunk training is a general training exercise that can be completed by individuals. To maintain the correct postures is extremely important for maximizing the training effect. In practice, it is difficult for an individual to grasp the posture during the training and achieve a correct posture. The previous work on training support systems usually requires a depth camera or multiple cameras, which are costly. Because the visual feedback to the user is normally 2D bone images, it is difficult to grasp the posture by showing only one viewpoint or skeletons. To solve these issues, this work aims to estimate pose using deep learning approaches and propose a posture support system for trunk training using a single RGB camera. The posture estimation is performed to generate a 3D human model from the camera image, and the target posture and the current posture are superimposed and displayed in the proposed user interface. To clarify the difference between the target and the current postures of 3D models, we adopt markers for visualization at ten body parts (both hands and elbows, shoulder, knee, ankle) using color changes. The proposed support interface visualizes the posture correction and provides visual feedbacks to the user from two viewpoints.

1. はじめに

近年、深層学習を用いて、ひとつの画像から姿勢を推定する研究が多く行われている。本研究では、深層学習を利用して姿勢を推定する研究[7,9]を応用し、カメラ画像を用いた体幹トレーニングの姿勢支援手法の提案を行う。

体幹とは、身体の四肢と頭部を除いた重量の約48%を占める非常に大きな部位であり、四肢間の運動連結やバランスに関して重要な役割を果たしている。そのため、体幹筋の働きを重視したトレーニング方法をスポーツ選手が行っており、その重要性が一般の方にも注目されるようになっている。そして、この体幹を鍛える体幹トレーニングは、

Worldwide Survey of Fitness Trends のフィットネス流行予測において、2007年から2010年まで、トップ5にランクインしており、2010年以降もずっとランクインしている[1]。このことから、体幹トレーニングが世界的に注目されていると言える。

体幹トレーニングは一般的に個人で行うことができ、体幹筋を中心とした全身の筋肉を鍛えることができる実用的かつ簡単なトレーニング手法である。そして、その効果を最大限に発揮するには、正しい姿勢を保つことが極めて重要である[2]。正しくない姿勢で行うトレーニングは怪我につながる可能性がある。しかしながら、個人がトレーニング中の姿勢を把握するのは難しい。そのため、鏡やカメラで撮影した映像を通じて姿勢を把握する方法があるが、把握した姿勢が正しいかどうかを判断することも難しい。このことから、体幹トレーニングでの姿勢支援は有効であると言える。

^{†1} 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa 923-1292 Japan
a) s1710236@jaist.ac.jp
b) xie@jaist.ac.jp
c) miyata@jaist.ac.jp



図 1 UI

トレーニング支援手法として、姿勢推定を行い、ユーザにフィードバックし、正しい動作、姿勢でのトレーニングを促す研究[4-6]がある。それらの姿勢推定には、深度カメラや複数台のカメラを必要とするため、コストがかかる。また、ユーザへの視覚的フィードバックが、1 視点および骨格情報のみを示している 2 次元ボーン画像であるため、姿勢の把握が難しい。

そこで本研究では、単一 RGB カメラのみを用いて、カメラ画像から姿勢推定を行う。そして、推定結果に基づいて 3 次元モデルを生成し、2 視点でユーザへ視覚的フィードバックをすることで、体幹トレーニングの姿勢支援を行う。2 視点である理由は、3 次元空間の情報を複数視点で提供することでユーザの空間把握を促進する研究[3]があり、姿勢の把握にも複数視点で行うことで、把握しやすさが向上するのではないかと考えたからである。しかしながら、視点数が多すぎると情報過多になってしまう可能性があるため 2 視点とする。

2. 関連研究

2.1 深度カメラを用いたトレーニング支援

高久ら[4]は筋力トレーニングの中でも比較的簡易で認知度の高い腹筋運動（シットアップ）を対象に支援を行っている。この研究は、深度情報とカラー画像を同時に取得できる Kinect を用いてユーザの筋力トレーニングの動作を検出する。そして、検出したトレーニング動作を速度、角度の観点から支援を行っている。また、ユーザのモチベーションを高めるため、トレーニングによる効果やゲーム性を付加している。見本動作の提示には、骨格情報のみを示しているボーン画像を用いており、付加情報として、推定消費カロリーを表示している。

また、岡本ら[5]は陸上競技者を対象としたハードルまたぎ練習の支援システムを提案している。この研究では、深度カメラである Kinect で取得した深度情報をもとにユーザの関節位置を推定し、ユーザの姿勢を取得する。そして、取得した姿勢をもとに適切かどうかを判断し、その結果からフィードバック情報を提示する。そのフィードバック情

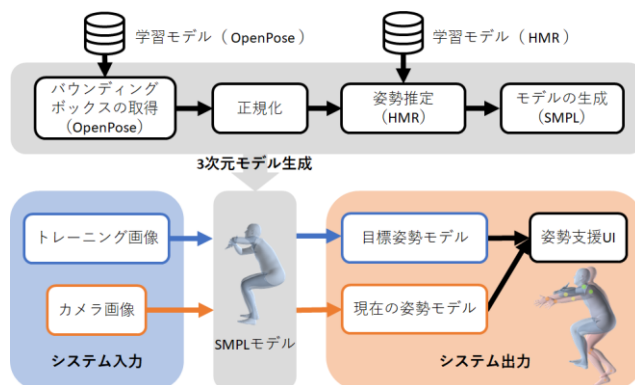


図 2 フレームワーク

報は、撮影している映像に重畳表示させ、文字やマーカなどで情報を示している。このシステムでは、マーカは正しい姿勢、誤った姿勢、訓練上目標となっている位置をそれぞれ指定している色で示しており、文字は事前に支援システムのプロトタイプを作成し、体験した陸上競技者のアンケート調査から数単語に限定して表示している。

2.2 複数のカメラを用いたトレーニング支援

松村ら[6]は、2 台の Playstation®Eye を用いて作成した 2 眼カメラを使って、弓道の練習を対象としたフォーム改善システムを開発した。この研究では、事前にお手本動画を取得し、半透明の動画にする。さらに、リアルタイムで撮影した映像と半透明にしたお手本動画を重畳した映像を Head Mounted Display に表示する。

2.3 本研究の位置づけ

以上で述べたトレーニング支援の関連研究では、深度カメラや複数台のカメラを用いてトレーニング支援を行っている。しかしながら、それらはコストがかかるため、本研究では、汎用性が高い単一の RGB カメラのみを用いて支援を行う。提案するシステムは、事前トレーニング画像からトレーニング姿勢を推定する。この手法の利点は、参考書やインターネットなどの画像でも、トレーニングの姿勢を取得できることである。すなわち、多種目への応用が容易にできる。トレーニング中の姿勢のフィードバックに関して、既存研究では、ボーン画像や撮影した映像への注釈、お手本動画と撮影した映像の重畳などであり、ユーザの視点が固定されている。ユーザが見たい視点からの映像を提供することにより、姿勢を把握しやすくなるのではと考え、本研究では、カメラ画像から姿勢推定を行い、3 次元モデルを生成し、ユーザが見たい視点で見ることが出来る映像を 2 視点で提供することとした。

本研究で作成したユーザインターフェース (UI) の例を図 1、フレームワークを図 2 に示す。フレームワークの 3 次元モデル生成は 3 節、システム全体の流れおよび UI は 4 節で説明する。

3. 画像からの3次元モデルの生成

3.1 手法

画像からの3次元モデル生成に、本システムでは人物像のバウンディングボックスの取得に OpenPose, 3次元モデルの生成に SMPL (A Skinned Multi-Person Linear Model), 姿勢推定に HMR (Human Mesh Recovery) を用いた。

OpenPose とは, Cao ら[7]が提案した画像から複数人の姿勢を推定する手法である。この手法は, multi-stage CNN を用いて, 入力画像から身体の一部が存在する位置と身体の一部間のつながり得る可能性を表すベクトルを求め, 姿勢を推定する。

SMPL とは, Loper ら[8]が提案した人間の様々な姿勢や体形を表現することができる人体モデルである。SMPL は, $N = 6,980$ の頂点をもつ三角メッシュ $M(\theta, \beta) \in \mathbb{R}^{3 \times N}$ で表現する。Pose $\theta \in \mathbb{R}^{3 \times K+3}$ は3次元の回転であり, 関節 $K = 23$ の軸角度と全体の回転で表す。

すなわち, θ は $3 \times K + 3 = 3 \times 23 + 3 = 72$ 個のパラメータを有している。Shape $\beta \in \mathbb{R}^{10}$ は, 主成分分析を用いて 10 次元で表された形状空間のパラメータである。図 3 に頂点, メッシュ, θ , β を代入した SMPL を示す。

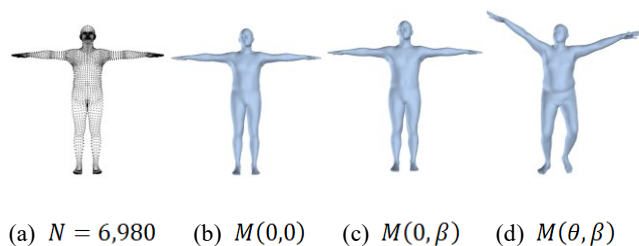


図 3 SMPL

HMR とは, Kanazawa ら[9]が提案した, 単一の RGB 画像から人体モデルを再構成する手法である。画像から 3D 回帰モジュールによって推測されたパラメータ (カメラの回転, 移動, スケールの 3 つと SMPL のパラメータである Pose θ , Shape β) で生成された人体モデル (SMPL) を不自然な関節の曲がり方をしない角度制限や, 細すぎる体など不自然な体型にしない制限などを学習した識別器によって自然な人体モデルであるかを判別する。自然なモデルであると判断されたモデルは出力され, 違うと判断された場合は, もう一度 3D 回帰モジュールによってパラメータを推測する。

3.2 3次元モデル生成

本システムでは, 画像に対して姿勢推定を行う前に, 画像の正規化を行う。正規化を行う理由は, 姿勢推定で用いる HMR の学習データが以下の 2 つの条件に統一されているためである。

- 1) 人物のバウンディングボックスの対角がおおよそ 150pixel
- 2) 縦横それぞれ 224pixel

正規化を行う際, 人物のバウンディングボックスを用いる。バウンディングボックスの取得には, OpenPose で得られる各関節の位置座標を用いる。

上の条件に合うように正規化した画像に対し, HMR を用いて姿勢推定を行う。4.2 節で後述するユーザへの視覚的フィードバックとして, 目標姿勢と現在の姿勢を重畳表示するため, 3次元モデルの形状や大きさを合わせる必要がある。そのため, 各関節の軸角度および全体の回転を表す Pose θ のみを SMPL に代入し, 3次元モデルを生成する。

4. 姿勢支援手法

4.1 概要

本システム全体の流れ (図 2) として, トレーニング前のトレーニング画像, トレーニング中のカメラ画像を入力として 3次元モデルを生成する処理は同様であるが, 以降の処理が異なる。トレーニング前は視点と 3次元モデルの保存, トレーニング中は, 視覚的フィードバックの生成を行う。以降, トレーニング前, トレーニング中の処理について詳述する。

まず, トレーニング前の処理について説明する。ユーザは, 行いたいトレーニングの画像をはじめに入力する。システムは, 入力した画像に対して 3次元モデルを生成する。そして, ユーザは生成された 3次元モデルの姿勢が分かりやすい視点の設定を行う。その視点と生成した 3次元モデルは, トレーニング中の処理で使用するため, 保存する。

次に, トレーニング中の処理について説明する。トレーニング中は, Web カメラからユーザのトレーニング画像を取得し, トレーニング前の処理と同様に, 3次元モデルの

生成を行う。そして, ユーザへの視覚的フィードバックとして, 生成された現在姿勢の 3次元モデルと事前に保存した目標姿勢の 3次元モデルを重畳表示する。また, 目標姿勢と現在の姿勢の違いを分かりやすくするために, 身体の一部 10 箇所 (両手および肘, 肩, 膝, 足首) にマーカを表示し, 目標姿勢との誤差に応じて, 色を 4 段階に変化させ, 姿勢の補正を促す。これらの情報は姿勢支援 UI を通じて, ユーザへフィードバックされることで, ユーザは姿勢を目標姿勢へと近づけていく。

以上の流れで, 本システムはユーザへの姿勢支援を行い, 正しい姿勢でのトレーニングを促す。

4.2 視覚フィードバックの生成

目標姿勢と現在姿勢の 3次元モデルを重畳させた画像は, トレーニング前の処理で保存された目標姿勢の 3次元モデルと現在姿勢の 3次元モデルを同時にレンダリングするこ

とで、重量表示する．2つの3次元モデルの原点は、腰部分で一致しているため、位置合わせは行わない．

次に、身体各部位 10箇所（両手および肘、肩、膝、足首）の誤差を示すマーカの表示手法について述べる．

まず、マーカの表示位置の選出理由として、末端の部位は体幹部分よりも誤差が大きくなりやすく、姿勢を補正するには重要な部分であると考えたためである．さらに、体幹部分の姿勢提示として股関節および肩の表示を考慮した．股関節は3次元モデルの原点が腰部分であることから、大きな誤差が出にくく、末端（足首、膝）の表示で十分と考え、本システムの姿勢支援には適切でないと考えた．肩に関しては、原点である腰部分から離れていることや、肘、手だけの表示では上半身の姿勢提示に不十分であると考え、肩を表示することとした．

次に、各部位のマーカの表示位置について述べる．マーカの表示位置は、関節位置を記した画像（[8] Figure 3(b)）と本システムで出力した SMPL の画像を重ね合わせ、各部位の画像座標の始点と終点（図 4）を取得する．そして、透視投影変換行列に本システムのレンダリング条件を代入した式(1)に始点と終点の大小関係を代入した式(2)を満たす3次元モデルのメッシュ頂点を探索した．メッシュ頂点は、視点によって表示位置がずれるのを防ぐために、前面と背面で1点ずつ取得し、表示にはその2点の平均値を用いる．図 5 に各部位に対応したメッシュ頂点の番号および位置を示す．色変化には、マーカの表示位置を求めるのに用いた3次元モデルのメッシュ座標の平均値を現在姿勢、目標姿勢とでそれぞれ求め、その2つの距離 d_{err} を用いる．条件は式(3)の通りである．

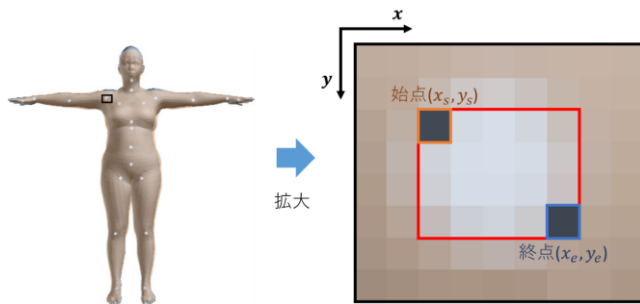


図 4 画像座標の始点と終点

$$\begin{bmatrix} x_{image} \\ y_{image} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{500 \times X_{model}}{Z_{model} + 2} + 332.5 \\ \frac{500 \times Y_{model}}{Z_{model} + 2} + 325 \\ 1 \end{bmatrix} \quad (1)$$

$$\begin{cases} \frac{x_s - 332.5}{500} (Z_{model} + 2) < X_{model} < \frac{x_e - 332.5}{500} (Z_{model} + 2) \\ \frac{y_s - 332.5}{500} (Z_{model} + 2) < Y_{model} < \frac{y_e - 332.5}{500} (Z_{model} + 2) \end{cases} \quad (2)$$

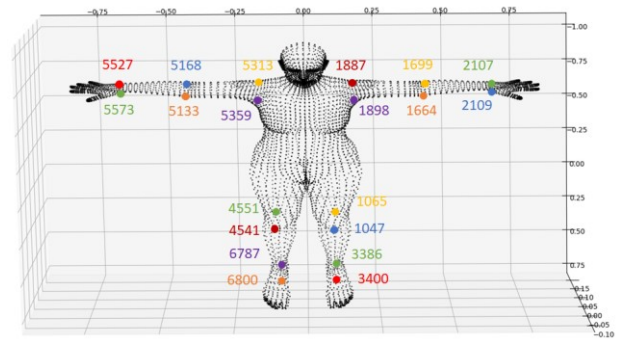


図 5 各部位に対応したメッシュ頂点番号

$$Color = \begin{cases} \text{red} & (\text{if } 0.5 \leq d_{err}) \\ \text{orange} & (\text{if } 0.25 \leq d_{err} < 0.5) \\ \text{yellow} & (\text{if } 0.1 \leq d_{err} < 0.25) \\ \text{greenyellow} & (\text{if } d_{err} < 0.1) \end{cases} \quad (3)$$

4.3 ユーザインターフェース

本研究では、図 6、図 7 のような UI を作成した．

まず、図 6 の画面では、トレーニングしたい画像を選択し、ユーザは見やすい視点を設定する．表示されている目標姿勢モデルをマウスでドラッグすることにより視点が回転する．視点を設定し終わると、図 7 の画面に切り替わる．

図 7 の画面では、カメラ画像と入力画像のウィンドウを表示することができる．そして、トレーニング中になると、現在の姿勢モデルと目標姿勢モデルを重量表示し、誤差を示すマーカも表示する．

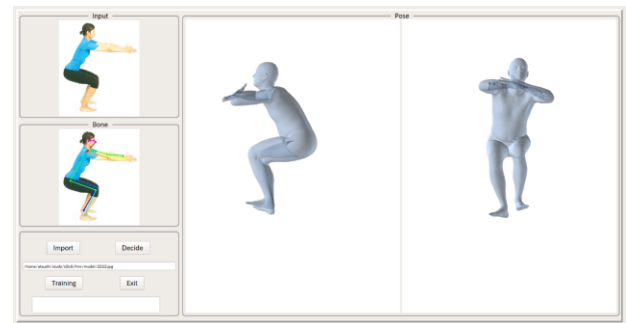


図 6 トレーニング前の UI

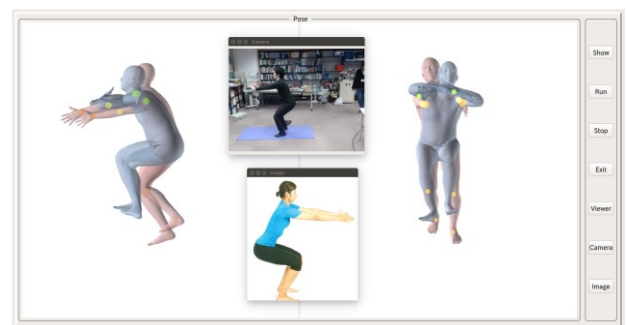


図 7 トレーニング中の UI

5. 評価実験

5.1 概要

本研究で提案した手法の有用性を検証するために評価実験を行う。関連研究[4]でフィードバック情報として用いている2次元ボーンを1視点で提示する手法(手法1)と、本研究で提案した3次元モデルを2視点から提示する手法(手法2)を比較する。手法ごとに異なる推定手法を用いると提示する目標姿勢に誤差が生じる可能性があるため、本研究で用いた推定手法を利用して、手法1を実装した。

提示する関節数は、OpenPose で用いられている COCO データセット[10]を参考に 18 箇所(両耳、両目、鼻、首、両肩、両肘、両手、両腰、両膝、両足首)とする。手法1、手法2のUIおよび提示例を図8、図9に示す。

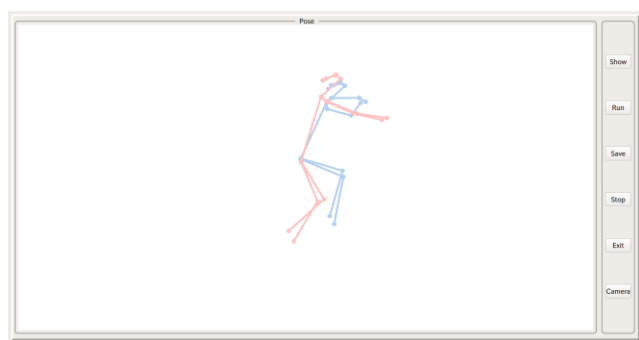


図8 UIおよび提示例(手法1)

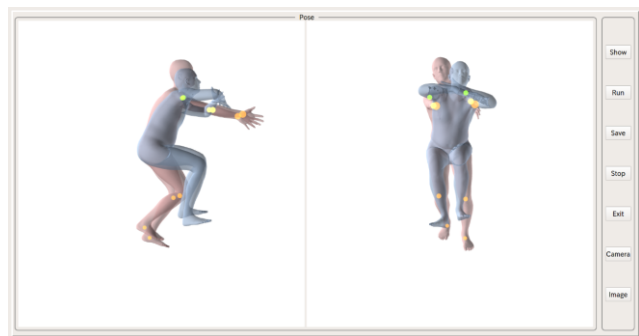


図9 UIおよび提示例(手法2)

実験の対象とする体幹トレーニングは図10(a)～図13(a)に示す4つであり、体位による影響を考え、立位、側臥位、仰臥位で行うトレーニングを参考文献[11]から選定した。対象とする体幹トレーニングそれぞれにおいて、手法1、手法2で提示するボーンおよび3次元モデル画像を図10(b), (c)～図13(b), (c)に示す。

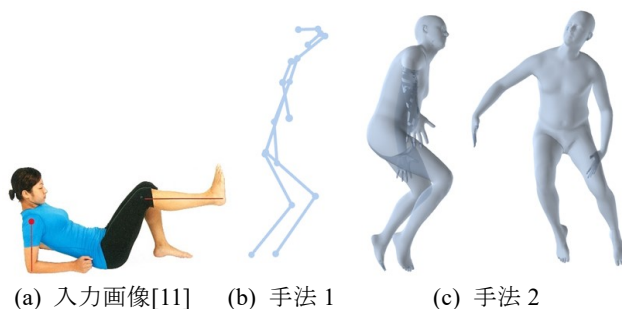


図10 体幹トレーニング1

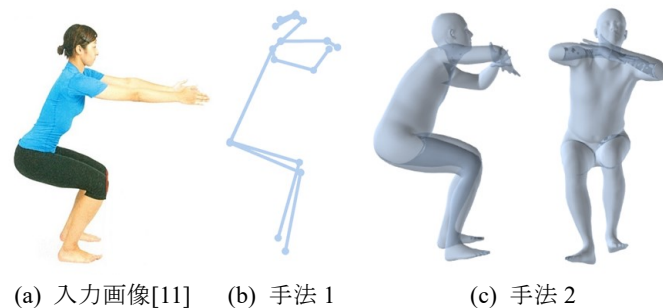


図11 体幹トレーニング2

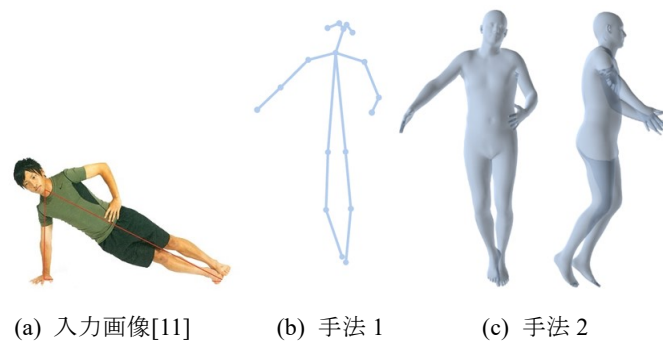


図12 体幹トレーニング3

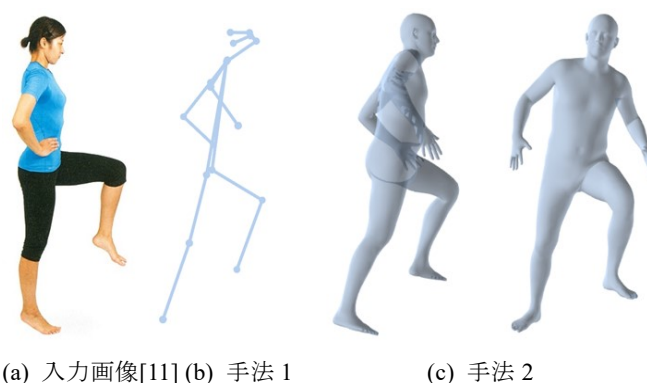


図13 体幹トレーニング4

本実験の被験者は、本学の男子学生8名(日本人学生:5名, 中国人留学生:3名, 年齢: 24.2 ± 1.05)である。手法1, 2と体幹トレーニングの組み合わせは、実施する順番による影響を考慮し、表1のようにする。体幹トレーニングはそれぞれ25秒程度行い、各フレームにおけるカメラ画

像およびトレーニング姿勢を保存する。また、実験終了後にアンケート（表 2）を実施し、姿勢の把握しやすさや本システムで付加した視覚的フィードバックについて調査する。

表 1 手法と体幹トレーニングの組み合わせ

	A	B	D	C	E	F	G	H
1 回目	2-1	2-3	1-4	1-2	2-1	2-3	1-4	1-2
2 回目	2-2	2-4	1-3	1-1	2-2	2-4	1-3	1-1
3 回目	1-3	1-1	2-2	2-4	1-3	1-1	2-2	2-4
4 回目	1-4	1-2	2-1	2-3	1-4	1-2	2-1	2-3

（手法番号 - 体幹トレーニング番号）

表 2 アンケート内容

No	設問	回答
1	これまでに体幹トレーニングをしたことはありますか	はい・いいえ
2	どちらの条件が姿勢を把握しやすかったですか	1 視点ボーン画像・2 視点 3 次元モデル
3	2 視点になることで、姿勢の把握がしやすくなりましたか	5・4・3・2・1 しやすい⇔全く
4	視点が変えられることで、姿勢の把握がしやすくなりましたか	5・4・3・2・1 しやすい⇔全く
5	3 次元モデルになることで、姿勢の把握がしやすくなりましたか	5・4・3・2・1 しやすい⇔全く
6	色マークは姿勢の補正に役立ちましたか	5・4・3・2・1 しやすい⇔全く
7	マークの大きさが近いほど大きく、遠いほど小さくなることに気が付きましたか	気が付いた・気が付かなかった
7-1	大きさの変更で見やすくなりましたか (7 の回答：気が付いた方のみ)	5・4・3・2・1 見やすい⇔全く
8	システムが提示したモデルとあなたが思っている姿勢はどの程度一致しましたか	5・4・3・2・1 思った通り⇔全く

本実験の評価指標として、平均平方二乗誤差（RMSE: Root Mean Square Error）を用いる。RMSE は式(4)で表される。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\vec{E}_i - \vec{M}_i)^2} \quad (4)$$

n は 6,980, $\vec{E}_i$ は被験者におけるトレーニング姿勢の 3 次元モデル座標ベクトル $\vec{E}_i = (x_{ei}, y_{ei}, z_{ei})$ である。 $\vec{M}_i$ は目標姿勢の 3 次元モデル座標ベクトル $\vec{M}_i = (x_{mi}, y_{mi}, z_{mi})$ であ

る。図 14 に示すように被験者は UI が表示されている画面を見ながら姿勢の補正を行う。そのため、顔の付近では目標姿勢と被験者のトレーニング姿勢の 3 次元モデルが一致しない。そこで、首より上のメッシュ頂点を除いて、RMSE を求めることとした。



図 14 実験風景の例

5.2 実験結果

実験結果を図 15～図 22, アンケート結果を表 3 に示す。

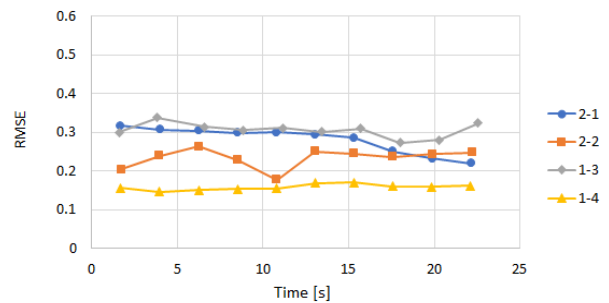


図 15 被験者 A

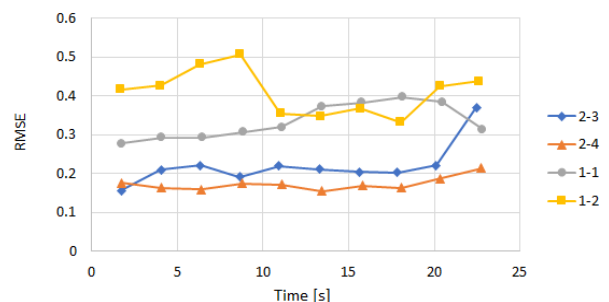


図 16 被験者 B

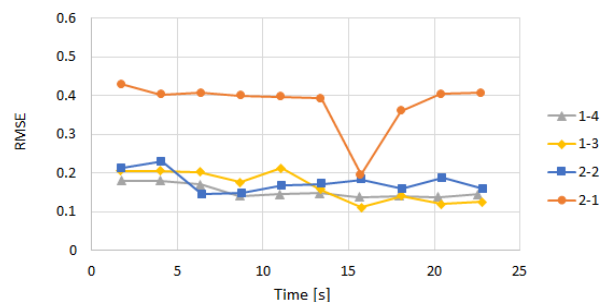


図 17 被験者 C

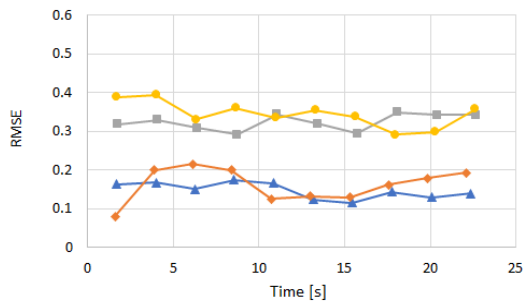


図 18 被験者 D

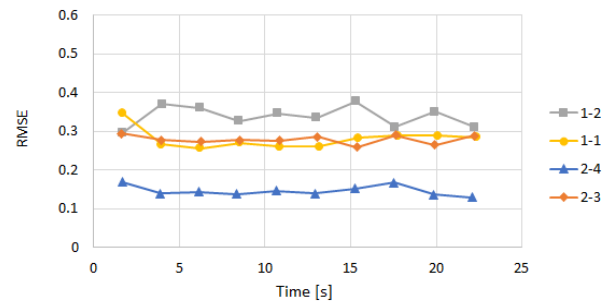


図 22 被験者 H

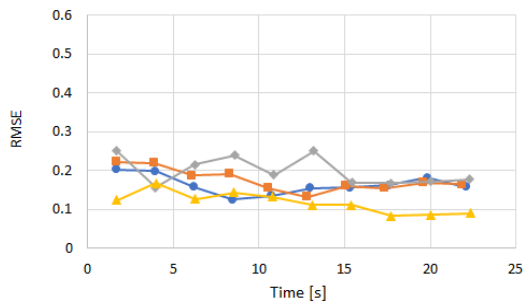


図 19 被験者 E

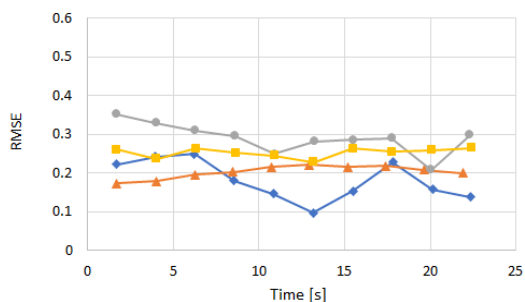


図 20 被験者 F

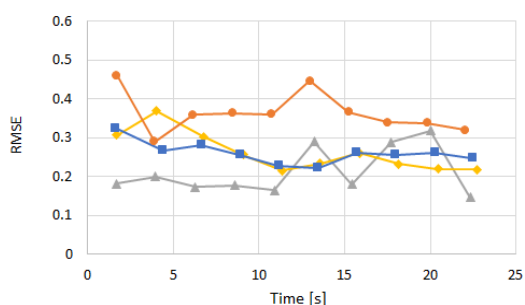


図 21 被験者 G

表 3 アンケート結果

被験者	設問								
	1	2	3	4	5	6	7	7-1	8
A	○	2	3	4	4	2	×	-	4
B	×	2	3	4	5	2	×	-	3
C	○	2	2	4	5	4	×	-	4
D	○	2	5	5	5	5	○	5	5
E	×	2	5	5	5	5	×	-	5
F	○	2	5	3	5	4	×	-	3
G	○	2	4	5	4	3	×	-	3
H	×	2	3	5	5	3	×	-	4

(設問 1「○：はい，×：いいえ」設問 2「1：1 視点ボーン画像，2：2 視点 3 次元モデル」設問 7「○：気が付いた，×：気が付かなかった」)

5.3 考察

はじめに，実験結果について考察をする．体幹トレーニング 3，4 は手法を問わず，RMSE が小さい．すなわち，容易なトレーニング姿勢であると言える．RMSE が急激に減少した直後の RMSE 値は大きくなっていることが多い．これは視覚フィードバックの更新速度が 2 秒程度の低速であることが起因していると考えられる．

姿勢補正への影響を調べるために，RMSE の最小値と初期値の差の初期値を基準とした割合の平均値および最小値に達する時間の平均値を示す．ここで，時間の平均値を求める際に 0 は除いた．表 4 より，手法 1 よりも手法 2 の方が 7% 程度大きいことが分かる．このことから，手法 2 の方が姿勢補正への影響が大きいと推測する．また，最小値に達する時間の平均値には大きな違いがないと考えられる．

表 4 実験結果のまとめ

割合[%]		時間[s]	
手法 1	手法 2	手法 1	手法 2
33.88	40.18	11.77	11.93

次に、アンケート結果について考察する。表 5 に得られたアンケート結果の平均値を示す。被験者の全員が 2 視点 3 次元モデルを姿勢の把握がしやすいと回答していることから、本研究で提案した手法は有効であると推測する。

アンケート結果の平均値が最も低かった回答は設問 6 の“色マークは姿勢の補正に役立ちましたか”であり、被験者 8 名のうち、2 名が“2”と回答している。さらに、設問 7 の“マークの大きさが近いほど大きく、遠いほど小さくなることに気が付きましたか”では、1 名を除く 7 名が“気が付かなかった”と回答している。このことから、被験者は実験中マークなどを確認する余裕はなく、提案手法では 3 次元モデルが提示するため、モデルの姿勢を真似することで姿勢の補正を行っていると考ええる。すなわち、何度も繰り返し本システムを利用し慣れることで、マークなどを見る余裕も生じ、より効率的な姿勢の補正が行えるのではないかと推測する。

設問 3 の“2 視点になることで、姿勢の把握がしやすくなりましたか”について、設問 6 に次いで結果が低かった。2 視点になることで姿勢の把握がしやすくなるのではないかと考え実装したが、本実験ではよい結果を得ることは出来なかった。一方、設問 4、5 の結果は高かったため、視点変更や 3 次元モデルを用いて視覚的フィードバックを行うことは姿勢の把握しやすさを向上するのに、有効な手法であると考ええる。

表 5 アンケート結果 (平均値)

設問					
3	4	5	6	7-1	8
3.75	4.38	4.75	3.50	5.00 (回答者 1 人)	3.88

6. おわりに

本研究では、カメラ画像を用いた体幹トレーニング支援手法の提案およびシステムを構築した。はじめに、深層学習を利用した研究を応用し、カメラ画像から人体 3 次元モデルを生成した。次に、その生成した 3 次元モデルを用いて、トレーニングの目標姿勢と現在の姿勢の違いを把握しやすくするための視覚的フィードバックの生成を行い、体幹トレーニングの支援システムを構築した。

関連研究でフィードバック情報として用いられている 2 次元ボーンを 1 視点で提示する手法と本研究で提案した 3 次元モデルを 2 視点から提示する手法を本システムで実装し、比較する評価実験を行った。目標姿勢の 3 次元モデルと推定したトレーニング姿勢の 3 次元モデルの平方平均二乗誤差 RMSE を用いて比較を行ったが、大きな違いは見られなかった。しかしながら、同時に実施したアンケート調査によって、提案した手法の有効性があると考えられる。

本システムの問題点として、視覚的フィードバックの更

新速度の遅さが挙げられる。現状では約 2 秒に一回更新される。実際に利用する際に、フィードバックがすぐに返ってこないため、その時点での目標姿勢との誤差が分からず修正してしまい、目標姿勢から遠ざかってしまう可能性がある。そのため、処理の高速化が望まれる。また、トレーニング画像から生成した目標姿勢の 3 次元モデルは、入力画像と一致しているとは言い難い。より良い支援を行うには、推定精度の向上も望まれる。

そして、本システムを利用する際、必ずディスプレイを見なければならないという問題もある。この問題に対しては、HMD (Head Mounted Display) を利用することで解決できると考える。

また、本研究で評価指標に RMSE を用いたが、精確ではない可能性があるため、今後検討する必要がある。

参考文献

- [1] T. R. Walter, WORLDWIDE SURVEY OF FITNESS TRENDS FOR 2018: The CREP Edition, *ACSM's Health & Fitness Journal*, 21(6) pages 10-19, 2017.
- [2] J. Stephenson, A. M. Swank, Core Training: Designing a Program for Anyone, *Strength and Conditioning Journal*, 26(6) pages 34-37, 2004.
- [3] 梅川香澄, 井上智雄, 岡田謙一, サッカーの戦略会議を支援する複数視点を用いた協調作業空間, 情報処理学会論文誌, 1(1) pages 19-26, 2013.
- [4] 高久大輔, 中島克人, Kinect を用いた筋力トレーニング支援システム, 情報処理学会第 77 回全国大会, pages 2-437--2-438, 2015.
- [5] 岡本勝, 磯村智将, 松原行宏, 姿勢推定手法を活用したリアルタイム運動訓練支援環境, 人工知能学会第 30 回全国大会, 1C4-OS-13a-1, 2016.
- [6] 松村海沙, 小池崇文, Head Mounted Display を用いた三人称視点によるフォーム改善システム, 情報処理学会第 78 回全国大会, pages 4-357--4-358, 2016.
- [7] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, Y. Sheikh, OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, *arXiv preprint arXiv:1812.08008*, 2018.
- [8] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, M. J. Black, SMPL: A Skinned Multi-Person Linear Model, *ACM Trans. Graphics (Proc. SIGGRAPH Asia)*, 34(6) pages 248:1--248:16, 2015.
- [9] A. Kanazawa, M. J. Black, D. W. Jacobs, J. Malik, End-to-end Recovery of Human Shape and Pose, *Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [10] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C. L. Zitnick, Microsoft COCO: Common Objects in Context, *European Conference on Computer Vision (ECCV)*, pages 740-755, 2014.
- [11] 木場克己, 体幹力を上げるコアトレーニング, 成美堂出版, 2012.