

令和 6 年 6 月 4 日現在

機関番号：16101

研究種目：基盤研究(C)（一般）

研究期間：2021～2023

課題番号：21K12175

研究課題名（和文）実技学習支援を目的とした深層学習による3Dボディ生成システムの開発

研究課題名（英文）Development of a 3D pose estimation system using deep learning to support practical learning

研究代表者

獅々堀 正幹（SHISHIBORI, Masami）

徳島大学・大学院社会産業理工学研究部（理工学域）・教授

研究者番号：50274262

交付決定額（研究期間全体）：（直接経費） 3,100,000円

研究成果の概要（和文）：本研究では、画像や映像などに写っている2次元の画像から、被写体の3次元骨格情報を深層学習技術を用いて推定する技術を開発することを目的とする。従来、画像中の被写体の2次元骨格情報の推定は高い性能を発揮しているため、3次元骨格情報の推定は実用化レベルの精度が得られていない。そこで、本研究では同一人物では姿勢を変更しても骨格間の距離が変化しない点に着目した新しい推定手法を考案した。評価実験として、Human3.6Mデータセットを用いた実験より、全体的な平均誤差を改善することができた。また、推定骨格の関節間距離が正解骨格に近いほど効果が高まることを確認した。

研究成果の学術的意義や社会的意義

実技を伴う教科では、実際の姿勢や動作、手技の動きなどを見せるために、事前に撮影した画像や映像をタブレット端末に配信し、実技教育を実践する教育現場が増えている。画像や映像を利用した実技教育は効率的に学習できるが、画像や映像を撮影した一方向からの視点でしか被写体を把握できず、身体に隠れた部分や視点を変更した内容が確認できないといった問題点がある。本研究成果を用いると、被写体の骨格情報を3次元で立体的に表示できるので、学習者は様々な視点から被写体を確認することで、学習効果が向上することが期待できる。

研究成果の概要（英文）：The technology for acquiring human pose is used in a wide range of fields such as motion analysis, AR, and VR. 2D pose estimation is to identify the 2D points of a person's joints in an image. Currently, 2D pose estimation is showing high performance. For this reason, 3D pose estimation has been widely studied using the results of 2D pose estimation as input. 3D pose estimation acquires depth information at joint points. The whole body or a part of the body leans forward or backward more than actual. In this study, we consider the fact that distances between joints of the same person is constant. By using this feature, this method improves the accuracy of the 3D pose estimation. In order to evaluate this method, we compared the 3D joint point positions estimated from the 2D joint point positions in the Human3.6M dataset. From the experimental result, it has been found that the more correctly distances between joints in estimated, the smaller the error tends to be.

研究分野：教育工学

キーワード：学習支援 拡張現実 深層学習 姿勢推定

## 様式 C-19、F-19-1 (共通)

### 1. 研究開始当初の背景

画像や映像を利用した実技教育は効率的に学習できるが、画像や映像を撮影した一方向からの視点でしか被写体を把握できず、身体に隠れた部分や視点を変更した内容が確認できないといった問題点がある。この問題に対処するため、画像や映像に写っている被写体を立体的に表示すると、 $360^\circ$  様々な視点から被写体を確認することが可能になり、学習効果が向上することが期待できる。しかし、2D 被写体を 3D モデルに変換するためには、200 万円以上する 3D ボディスキャナーが必要になる。また、数百枚の画像から 3D モデルを作成するソフトウェアも存在するが、処理時間と精度に問題が残る。本研究では、このような特殊な機材を必要とせずに、通常の環境で撮影された画像や映像に写っている被写体を 3D ボディに変換する。

### 2. 研究の目的

本研究の目的は、深層学習技術を用いて、画像や映像などの平面状 (2D) に写っている人物を立体的な 3 次元モデルに変換する技術を開発することである。2 次元で撮影された人物画像の 3 次元骨格座標を推定するタスクを 3D 姿勢推定と呼ぶ。現在の 3D 姿勢推定モデルは関節点の奥行き情報を取得する際、実際より体全体や体の一部が前後に傾いたり、動きの大きい手足などの末端の位置に大きなずれを生じるといった問題がある。本研究では、既存手法として現在最も高い精度が得られる LCN (Locally Connected Network)[1]に着目し、LCN を上回る精度が得られる新しい 3D 姿勢推定手法を提案する。また、大規模なデータセットを用いた実験により、提案手法の有効性を検証することを研究目的とする。

### 3. 研究の方法

#### (1) 提案手法の概要

現在、2D 姿勢推定は高い性能を発揮している。このため、3 次元座標を特定する 3D 姿勢推定は、2D 姿勢推定の結果を入力とする研究が盛んである。3D 姿勢推定モデルの一つである LCN は、2D 関節点座標のみを入力として 3D 姿勢を推定するため、画像の外観情報によって学習されたモデルが汎化性能を損なう恐れがない。

本研究では、既存手法として LCN に着目し、被写体が同一人物の場合、どのような姿勢をしても関節間距離が一定であることを考慮することで、3D 姿勢推定法の精度向上を図る。画像や映像内から複雑でない姿勢を選び、その推定骨格から人物の関節間距離を計算し、基準となる骨格長を求める。基準の骨格に合わせるように他の姿勢に対して関節点の奥行き方向を決定する。これにより、関節点間の距離が変化する問題を改善する。評価実験として、Human3.6M[2]のデータセットにある 2D 関節点座標から推定した 3D 関節点座標位置と実際の 3D 関節点座標位置を比較した。実験結果より、全体的な平均誤差を改善することができ、基準となる姿勢における推定骨格の関節間距離が正解骨格に近いほど効果があることを確認できた。

#### (2) LCN の概要と問題点

既存手法として着目した LCN の構造について説明する。LCN は表現力と汎化能力といった二つの長所を併せ持つモデルであり、従来手法の中でも高い精度が得られることが知られている。LCN の重みパラメータは関節ごとに独立しており、また、近傍の関節との位置関係に従って個々の関節情報を推定できる。LCN が用いられるモデルの概略図を図 1 に示す。

図 1 は、二つ隣の関節までの座標情報を利用するように設定された LCN のモデル構想を示す。例えば、関節 A の座標情報は、A、B、C の関節の座標情報を利用して推定される。同様の処理を繰り返すことにより、重みパラメータは関節ごとに独立させたまま、畳み込み処理の効果を持たせることができる。この仕組みを実現するため、LCN は重みパラメータ行列の他に構造行列を導入している。構造行列とは、ある関節点の座標情報を推定する際に、どの関節点の座標情報を利用するかを行列で表現したものである。

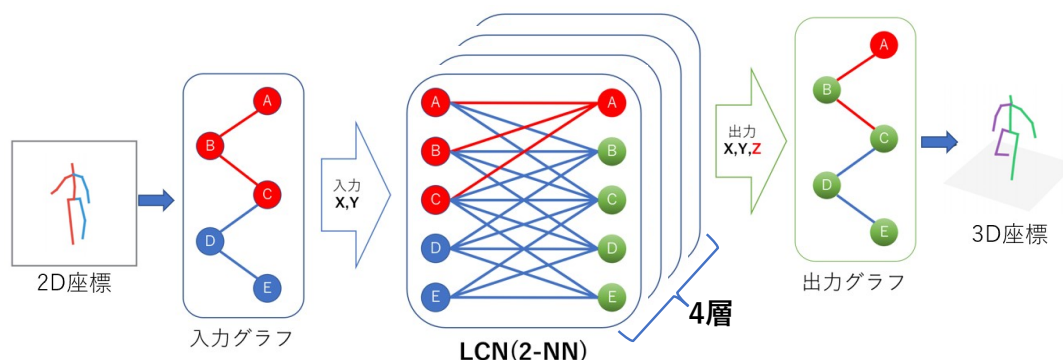
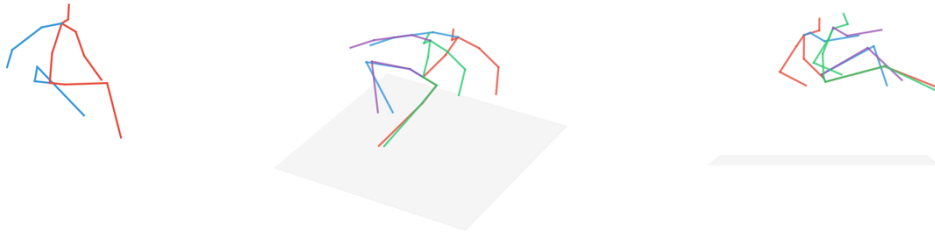


図 1 LCN の概略図

LCNにはK-NNと呼ばれる概念が存在する。これは、LCNが一つの関節点を推定するとき、どこまで隣接した関節の座標情報を考慮するかを決定するものである。K-NNのKには整数値が入り、K個隣の関節までの座標情報を考慮することを意味する。例として、2-NNの場合、2個隣の関節までの座標情報を考慮するため、右手の関節の座標情報を推定する際には、自身と右肘と右肩の関節の座標情報を考慮することになる。本研究では、LCNの論文内で最も性能が良かった3-NNを基準とする。

LCNの問題点として、2Dの関節点位置から奥行き情報を推定する難しさによって、不正確な3D姿勢が生成されることがある。入力に使用した実際の2D関節座標で人体グラフ構造を生成した結果を図2(a)、LCNを用いて推定された3D関節座標の人体グラフと実際の3D関節座標の人体グラフを比較した結果を図2(b)、側面から見た結果を図2(c)に示す。図2より、後ろにもたれたポーズをとると、奥行き情報が正確に再現されていない3D姿勢が作成される。以上より、腰を原点とした3D姿勢推定は手足などの末端の関節に向かうにつれて、奥行き情報が不正確になり、実際の人物の関節位置との誤差が大きくなる。これにより、実際よりも前後に傾いた姿勢や関節間距離が異なる姿勢が推定される場合がある。



(a) 実際の2D姿勢 (b) 実際の3D姿勢(赤青)と (c) 図2(b)を側面から見た結果推定された3D姿勢(緑紫)

図2 LCNでの推定結果の失敗例

### (3) 提案手法

人物の関節間距離を考慮して後処理を施す提案手法を述べる。本研究では、同一人物の骨格の長さがどのような姿勢や動作であっても同等であることに着目し、LCNで推定した骨格座標を修正することで精度改善を図る。はじめに、同じ人物・動作内で基準となるフレームを決定する。本研究では、どの一連動作のデータに対しても、各動作の最初のフレームを採用する。次に、基準フレームの推定骨格から人物の関節間距離を計算し、その骨格に合わせるように他のフレームに対して関節点の奥行き方向を決定する。

具体的な計算手順を説明する。人体の関節点数を $n$ とする。本実験で用いるデータセットは17点なので、 $n=17$ である。次に、始点となる関節点番号を格納したベクトル  $p \in \mathbb{R}^{n-1}$ 、終点となる関節点番号を格納したベクトル  $q \in \mathbb{R}^{n-1}$  を式(1)、(2)のように定義する。関節点番号と関節点位置の対応は、LCNの参考文献[1]に準拠する。

$$p = [0, 1, 2, 0, 4, 5, 0, 7, 8, 9, 8, 11, 12, 8, 14, 15] \quad (1)$$

$$q = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16] \quad (2)$$

推定する関節の3次元座標を  $y(x, y, z) \in \mathbb{R}^{3n}$  とする。添え字は関節点番号を示し、指定がなければ  $i$  と置く。例として、 $y_3$  は右足の推定3次元関節点座標を示す。

2点間  $y_p, y_q$  の推定骨格ベクトルを  $k(x, y, z) \in \mathbb{R}^{3(n-1)}$  とする。 $k_i$  は次式で表される。

$$k_i = y_{qi} - y_{pi} \quad (3)$$

基準フレームにおける  $k$  のユークリッド距離を  $d \in \mathbb{R}^{n-1}$  とする。 $d_i$  は次式で表される。

$$d_i = \sqrt{\hat{x}_i^2 + \hat{y}_i^2 + \hat{z}_i^2} \quad (4)$$

式(4)の両辺を2乗し、以下のように  $z_{qi}$  を求める。

$$\hat{z}_i^2 = d_i^2 - \hat{x}_i^2 - \hat{y}_i^2 \quad (5)$$

$$z_{qi} = z_{pi} \pm \sqrt{d_i^2 - \hat{x}_i^2 - \hat{y}_i^2} \quad (\because \hat{z}_i = z_{qi} - z_{pi}) \quad (6)$$

$z_{qi}$  を  $i=0$  から  $i=16$  まで順に更新し、推定座標を置換することで関節間距離を考慮することができる。

$d_i < \hat{x}_i^2 + \hat{y}_i^2$  の場合、 $z_{qi}$  はどの位置に動かしても関節間距離が基準より大きくなるため、提案手法では  $z_{qi} = z_{pi}$  とする。前の番号の関節との位置関係から  $z_{qi}$  を一通りに決める。提案手法を適用前の時点で、 $z_{qi} \geq z_{pi}$  の場合、

$$z_{qi} = z_{pi} + \sqrt{d_i^2 - \hat{x}_i^2 - \hat{y}_i^2} \quad (7)$$

$z_{qi} < z_{pi}$  の場合、

$$z_{qi} = z_{pi} - \sqrt{d_i^2 - \hat{x}_i^2 - \hat{y}_i^2} \quad (8)$$

以上の計算によって、各フレームの関節間距離を基準フレームに合わせることができる。提案手法における補正過程を図3に示す。図3は  $z_{qi}$  を  $i=3$  ( $q_3=4$  より、 $z_4$  は左腰) から  $i=5$  ( $q_5=6$  より、 $z_6$  は左足) まで順に更新している。

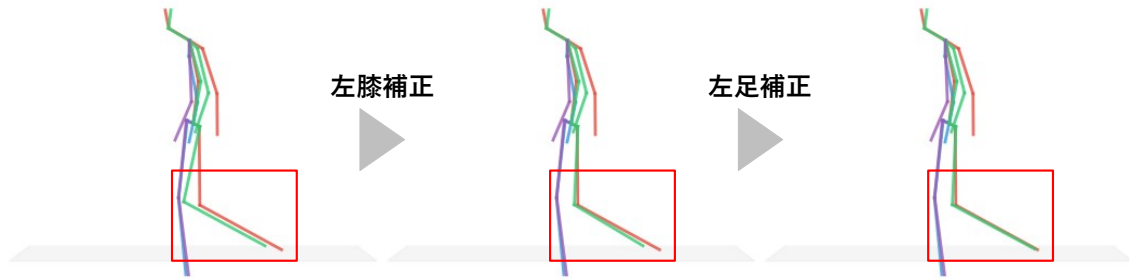


図3 補正過程のイメージ

#### 4. 研究成果

##### (1) 実験条件

###### ① データセット Human3.6M

Human3.6M とは、3D 姿勢推定モデルの学習、評価に標準的に用いられる大規模データセットである。データ数は約 360 万フレームあり、種類は動画や関節点の座標情報、3D メッシュなどがある。実験室環境でモーションキャプチャを用いてデータ作成しており、マーカーを装着した被験者を 4 台のモーションカメラで撮影することにより、各関節点の 3 次元座標をアノテーションしている。データは被験者別、シナリオ別に分けられている。本研究では、訓練データ 5 人、評価データ 2 人の被験者を用いてモデルの学習、評価を行う。各被験者は計 15 シナリオのそれぞれ異なった一連の動作を行っている。各シナリオは似た一連の動作 2 種類がある。この 2 種類の動画をサブシナリオ 1, 2 とする。そして、動画は 4 方向から撮影されているため、各被験者の各サブシナリオに 4 種類の動画が存在する。

###### ② 実験方法

評価実験では、Human3.6M の 5 人の訓練データを 200 epoch 学習させたモデルを用いる。また、訓練データとは別の Human3.6M の 2 人の評価データを用いて既存モデル (LCN) と提案モデルとの比較を行う。2D 関節座標と 3D 関節座標の正解値 ground-truth は、Human3.6M のモーションキャプチャで生成された実際の関節点位置データとする。

実験対象とするモデルは、以下の 4 種類とする。

- 既存のモデル LCN (以下、従来手法とする)
- 提案手法を従来手法に適用したモデル各動画の基準フレームをその動画内だけに適用 (以下、提案手法 1 とする)
- 提案手法を従来手法に適用したモデル各動画の基準フレームを被験者ごとに平均して適用 (以下、提案手法 2 とする)
- 提案手法を従来手法に適用したモデル基準フレームに各フレームの正解データを使用して適用 (以下、提案手法 3 とする)

基準フレームの数は提案手法 1 では 240、提案手法 2 では被験者ごとに平均するため、被験者数と同じ 2 フレームである。

### ③ 評価基準

評価指標として、MPJPE (Mean Per-Joint Position Error) を用いる。MPJPE は、3D 姿勢推定の性能を評価する上で最も広く使われている評価指標であり、全ての関節点ごとに推定された 3 次元関節座標と ground-truth(正解データ)座標の間のユークリッド距離の平均値となる。関節数を  $N$ 、 $J_i$  を  $i$  番目の正解座標位置、 $J_i^*$  を  $i$  番目の推定座標位置とする。計算式は次の通りになる。なお、Human3.6M のシナリオ別、関節点別に MPJPE を用いて評価を行う。

$$MPJPE = \frac{1}{N} \sum_{i=1}^N \|J_i - J_i^*\|_2 \quad (9)$$

### (2) 実験結果

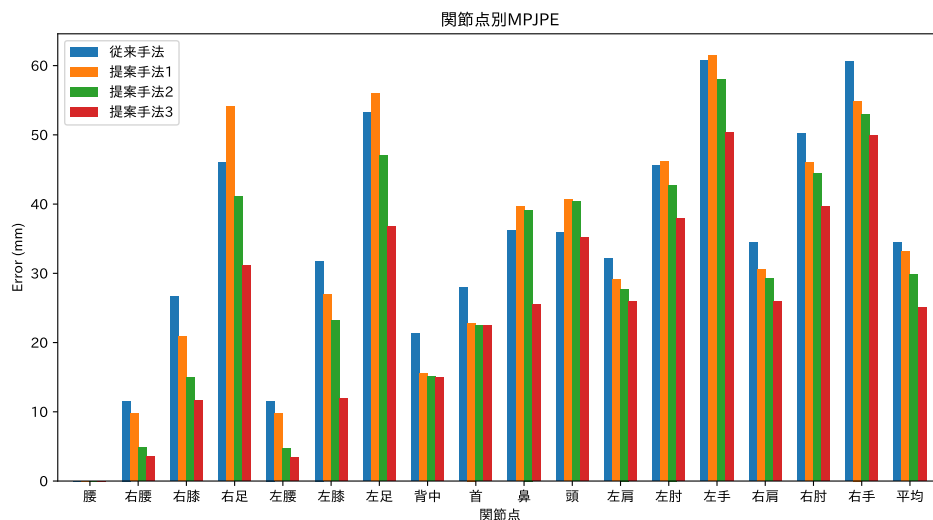


図 4 関節点別の MPJPE

関節点別の推定精度を図 4 に示す。従来手法と提案手法でシナリオ別 MPJPE と関節点別 MPJPE を計測した結果、全体的に従来手法よりも提案手法の方が誤差が小さくなり、性能が上がることを確認できた。以下、それぞれの提案手法についての特徴を分析する。

#### ● 提案手法 1

従来手法よりも精度は向上し、関節間距離を考慮することの有効性は実証された。また、主に原点である腰から比較的近い関節点で従来手法よりも誤差が小さくなった。腰周辺の関節間距離は正解に近いと考えられる。ただし、精度の改善は見られたが、正確な骨格間距離との誤差があるため、提案手法 2 及び 3 までの精度は得られていない。

#### ● 提案手法 2

提案手法 1 と比べて、総合的に誤差が小さくなっている。これは、複数のシナリオに登場する同一人物の被験者から関節間距離を得たことで、提案手法 1 よりも関節間距離が正解に近づいたためと考えられる。しかし、鼻と頭の MPJPE は従来手法よりも誤差が大きくなった。これは、鼻と頭の関節間距離は非常に短いため、提案手法の効果に影響しにくかったことが原因と考えられる。

#### ● 提案手法 3

正解骨格を基準としているため、提案手法 1 及び 2 よりも総合的に誤差が小さくなっている。今回の実験では、データセットから正解の関節間距離を取得したが、実用化に向けて、予め被験者の関節間距離を計測した後に画像や映像を撮影することができれば、正確な 3D 姿勢推定が可能になることが期待できる。

#### <引用文献>

- [1] H. Ci, X. Ma, C. Wang, Y. Wang, "Locally Connected Network for Monocular 3D Human Pose Estimation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.44, No.3, pp.1429-1442, March 2022, doi: 10.1109/TPAMI.2020.3019139.
- [2] C. Ionescu, D. Papava, V. Olaru and C. Sminchisescu, "Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments," in IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 36, No. 7, pp. 1325-1339, July 2014, doi: 10.1109/TPAMI.2013.248.

## 5. 主な発表論文等

〔雑誌論文〕 計5件（うち査読付論文 5件/うち国際共著 0件/うちオープンアクセス 5件）

1. 著者名 Mitsuhara Hiroyuki、Shishibori Masami	4. 巻 16
2. 論文標題 Tsunami Evacuation Drill System Focusing on Mobile Devices	5. 発行年 2022年
3. 雑誌名 International Journal of Interactive Mobile Technologies (iJIM)	6. 最初と最後の頁 4~20
掲載論文のDOI (デジタルオブジェクト識別子) 10.3991/ijim.v16i06.23419	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Liu Meiqin、Mitsuhara Hiroyuki、Shishibori Masami	4. 巻 16
2. 論文標題 A Mobile Application for Earthquake Education Targeting Foreigners Intending to Visit Japan	5. 発行年 2022年
3. 雑誌名 International Journal of Interactive Mobile Technologies (iJIM)	6. 最初と最後の頁 170~190
掲載論文のDOI (デジタルオブジェクト識別子) 10.3991/ijim.v16i24.35479	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Hiroyuki Mitsuhara and Masami Shishibori	4. 巻 1
2. 論文標題 Interactive Digital Signage System Enabling Viewers to Answer Quizzes by Body Actions	5. 発行年 2021年
3. 雑誌名 Information and Technology in Education and Learning	6. 最初と最後の頁 1-11
掲載論文のDOI (デジタルオブジェクト識別子) 10.12937/itel.1.1.Dev.p001	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Mitsuhara Hiroyuki、Tanimura Chie、Nemoto Junko、Shishibori Masami	4. 巻 7
2. 論文標題 Location-Based Game for Thought-Provoking Evacuation Training	5. 発行年 2023年
3. 雑誌名 Multimodal Technologies and Interaction	6. 最初と最後の頁 59~59
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/mti7060059	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Mitsuhara Hiroyuki、Shishibori Masami	4. 巻 19
2. 論文標題 Shelter GO: multiplayer location-based game to promote knowing shelters for emergency evacuation	5. 発行年 2023年
3. 雑誌名 Research and Practice in Technology Enhanced Learning	6. 最初と最後の頁 009 ~ 009
掲載論文のDOI (デジタルオブジェクト識別子) 10.58459/rptel.2024.19009	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計11件 (うち招待講演 1件 / うち国際学会 9件)

1. 発表者名 Kaito Oe, Itsuki Tanioka, Hiroyuki Mitsuhara and Masami Shishibori
2. 発表標題 Immersive Function for Allocating Disaster Situations for a VR-based Evacuation Training System
3. 学会等名 Proceedings of the 30th International Conference on Computers in Education (国際学会)
4. 発表年 2022年

1. 発表者名 Kaito Oe, Itsuki Tanioka, Hiroyuki Mitsuhara and Masami Shishibori
2. 発表標題 Prototype System of Evacuation Training in Metaverse
3. 学会等名 Proceedings of the 30th International Conference on Computers in Education (国際学会)
4. 発表年 2022年

1. 発表者名 Hiroyuki Mitsuhara and Masami Shishibori
2. 発表標題 Shelter GO: Multiplayer Location-based Game for Learning Evacuation
3. 学会等名 Proceedings of the 30th International Conference on Computers in Education (国際学会)
4. 発表年 2022年

1. 発表者名 Hiroyuki Mitsuahara, TANIOKA Itsuki and Masami Shishibori
2. 発表標題 Observing Evacuation Behaviours of Surprised Participants in Virtual Reality Earthquake Simulator
3. 学会等名 Proceedings of the 29th International Conference on Computers in Education (国際学会)
4. 発表年 2021年

1. 発表者名 Hiroyuki Mitsuahara, Chie Tanimura, Junko Nemoto and Masami Shishibori
2. 発表標題 Expressing Disaster Situations for Evacuation Training Using Markerless Augmented Reality
3. 学会等名 Proceedings of 25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (国際学会)
4. 発表年 2021年

1. 発表者名 Harumi Kadota, Shoma Fujisawa, Masaki Oono, Ryoma Morigaki, Nobuaki Yamamoto, Yasuhisa Kanematsu, Manabu Ishihara and Masami Shishibori
2. 発表標題 Development of a 3D visualization system for the cerebral aneurysm coil embolization
3. 学会等名 Proceedings of Fifteenth International Conference on Quality Control by Artificial Vision (国際学会)
4. 発表年 2021年

1. 発表者名 Ryota Nishio, Masaki Oono, Takaharu Goto, Takahiro Kishimoto and Masami Shishibori
2. 発表標題 A reconstruction method of 3D face model from front and side 2D face images using deep learning model
3. 学会等名 Proceedings of Fifteenth International Conference on Quality Control by Artificial Vision (国際学会)
4. 発表年 2021年



1. 発表者名 獅々堀正幹, 大野将樹, 泓田正雄, 寺田賢治, 緒方良輔, 岡久稔也
2. 発表標題 胸腹水濾過濃縮専用装置 M-CART への人工知能応用の可能性
3. 学会等名 第59回日本人工臓器学会大会 (招待講演)
4. 発表年 2021年

1. 発表者名 Koshiro Hori, Takeo Minamikawa, Yoshiki Terao, Masami Shishibori and Takeshi Yasui
2. 発表標題 Identification of spectral features for selective detection of peripheral nerves by support vector machine-based Raman spectral analysis
3. 学会等名 Proceedings of SPIE (国際学会)
4. 発表年 2023年

1. 発表者名 Kaito Ohe, Yusaku Ichino, Hiroyuki Mitsuhara and Masami Shishibori
2. 発表標題 Non-Player Characters for Evacuation Training in Metaverse: Preliminary Experiment
3. 学会等名 Proceedings of the 31st International Conference on Computers in Education (国際学会)
4. 発表年 2023年

1. 発表者名 巨島 弘基, 谷口 拓也, 河原田 理愛, 尾矢 剛志, 常山 幸一, 大野 将樹, 獅々堀 正幹
2. 発表標題 細胞診断システムにおける円形度に着目した領域分割の改善
3. 学会等名 2023年電気学会 電子・情報・システム部門大会
4. 発表年 2023年

〔図書〕 計1件

1. 著者名 北 研二、松本 和幸、吉田 稔、獅々堀 正幹、大野 将樹	4. 発行年 2021年
2. 出版社 科学情報出版	5. 総ページ数 184
3. 書名 - Pythonでデータサイエンス - AI・機械学習のためのデータ前処理 [ 実践編 ]	

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	光原 弘幸  (MITSUHARA Hiroyuki)  (90363134)	徳島大学・大学院社会産業理工学研究部(理工学域)・教授    (16101)	
研究分担者	大野 将樹  (OONO Masaki)  (90433739)	徳島大学・大学院社会産業理工学研究部(理工学域)・講師    (16101)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------