EVALUATION OF MOTION DIFFERENCE BETWEEN INDIVIDUAL PERSONS

Yasushi Mae[1*], Akihisa Nagata[1], Tomokazu Takahashi[1], Masato Suzuki[1], Seiji Aoyagi[1], and Yasuhiko Arai[1]

(Received September 8, 2023; accepted January 19, 2024)

Abstract

Examining differences in motion between individual persons involves predicting the three-dimensional (3D) future pose using an RGB-D camera. The 3D pose of a person is estimated based on depth data corresponding to the two-dimensional pose on the RGB image. Utilizing a neural network model, a sequence of 3D poses representing human motion is learned, and the future 3D pose is predicted from the sequential input of these 3D poses. The experiment evaluates the error in predicting the future 3D pose for individual persons. The evaluation involves basic motions like ‘standing’ and ‘sitting,’ which do not involve a change of location in everyday environments. The results highlight individual differences in motion and demonstrate the effectiveness of using personal motion data for learning and predicting the motion of individuals.

Key words: 3D human pose prediction; Individual difference; RGB-D camera

1 Introduction

Service robots are designed to support and enhance human daily activities, making personalization in user experience a crucial aspect. Achieving this involves having home service and environmental robots observe a person’s actions, identifying the needed assistance and determining the best time to provide support by predicting the person’s motion.

For a service robot to exhibit intelligent assistance, its visual capabilities must closely match those of humans. In recent years, deep neural networks (DNNs) have demonstrated superior abilities in object recognition tasks, surpassing human performance. Additionally, DNNs can robustly estimate human poses, further contributing to the robot’s understanding of human actions and facilitating personalized support.

Thus far, many two-dimensional (2D)[1-4] and three-dimensional (3D) human pose estimation methods[5-12] have been proposed, with many relying on the relationship between 2D human pose and 3D human pose data obtained from 3D motion capture. Additionally, the prediction of future human pose has been previously studied[13-17], often involving the learning of diverse motions from datasets comprising contributions from numerous individuals. Traditionally, these methods use datasets with data collected from various individuals to

[1] Department of Mechanical Engineering, Kansai University, Suita, Osaka 564-8680, Japan

[*] Correspondence to: Yasushi Mae, Department of Mechanical Engineering, Kansai University, Suita, Osaka 564-8680. E-mail: mae@kansai-u.ac.jp

ensure general applicability.

With the increasing popularity of RGB-D cameras, they have become valuable visual sensing devices in robotics and computer vision. Unlike standard cameras, RGB-D cameras provide both RGB images and real-time depth information about a scene. This technology offers a distinct advantage as it allows direct measurement of depth information without relying on general 3D human motion datasets obtained through 3D motion capture. In[18), for instance, a method utilizing an RGB-D camera, specifically, Kinect v2, was employed to predict a person's future body position 0.5 s ahead during a jumping motion based on a dataset derived from the jumping motions of 11 individuals.

However, the aforementioned methods do not account for the unique differences and features in individuals' motion. Human motion varies among people, as each person possesses distinct features even when performing the same categorized motion, such as 'standing' and 'sitting'. Our target application in the future involves personalized robot services tailored to persons in everyday environments. For effective personalized service robots assisting with human daily activities, it is advantageous to create and learn a personal motion dataset for predicting individual motion. Further, predicting future 3D pose and motion, considering individual differences and features, can enhance the capabilities of service robots, enabling them to initiate movement and actions earlier in support of their users. In the application, we envision that vision systems introduced in everyday environments, including service robots, will measure and learn individuals' motion to predict future motions for assistance. We expect that training data including a diverse range of motions will be obtained from each individual person over long-term observation, resulting in personalized motion predictions for each person.

We evaluate the difference in motion among individuals based on the 3D human poses measured and predicted using an RGB-D camera and neural network models. We use a method presented in[19) for 3D human pose prediction in future frames. In this approach, the 2D pose of a person in the RGB image is estimated using OpenPose[2), an image-based pose estimation technique, while the 3D pose is obtained from the depth measured by the RGB-D camera. The motion of the person is learned using a neural network model, and the 3D pose in the future frame is predicted from the past sequence of the 3D pose[19). In[19), given the anticipated personal use of the application, the RGB-D camera is positioned in a room with an everyday environment. The motions of a person are measured for training, and the motion predictions are performed in the same setup. Since 'standing' and 'sitting' motions are basic motions that do not involve a change of location in everyday environments, the prediction error for these motions is evaluated in[19). In this paper, we examine the difference between individuals for 'standing' and 'sitting' motions as the first step in motion prediction. The paper experimentally evaluates the error in predicting the 3D future pose for individuals, revealing differences in individuals' motion and the effectiveness of using personal motion data for learning and predicting individual motions.

2 Prediction of 3D Human Pose

2.1 Measurement of 3D Human Pose using RGB-D Camera

We use an RGB-D camera to measure human poses, capturing both RGB and depth images

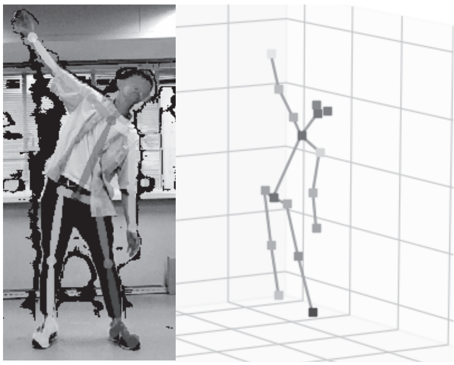


Figure 1. Examples of estimated 2D and 3D poses with RGB-D camera
Figure 1 is referred to in Section 2.1.

at the same time. The 3D pose of a person is measured following the approach described in [19]. First, the 2D pose of a person is estimated from the RGB image obtained by OpenPose[2], which detects the 2D skeleton consisting of several keypoints on the RGB image. An example of estimated 2D poses for a scene where ‘a person is exercising with motion’ is depicted in Fig. 1 on the left. The 2D skeletons are represented by connecting several keypoints, overlaying them on the input RGB image. The depth data is also measured by RGB-D cameras, and the captured RGB and depth images are aligned to match the image coordinates. Subsequently, the depth data corresponding to the 2D keypoints are obtained. Next, the RGB-D camera obtains a 3D pose as a set of 3D keypoints. The estimated 3D pose for the scene where ‘a person is exercising with motion’ is illustrated in Fig. 1 on the right, where the 3D skeleton is represented by 17 keypoints. In the example, we used the Intel RealSense D455 as the RGB-D camera.

Assuming that the 3D coordinates of a keypoint at time $\boldsymbol{t_n}$ are denoted by $\boldsymbol{P}(t_n)$, a 3D pose $\boldsymbol{Pose}(t_n)$ is represented by a set of keypoints $\boldsymbol{P}(t_n)$. The 3D motion is represented by a time sequence $\boldsymbol{Pose}(t_n)$. If depth data are not available at the pixel corresponding to the 2D keypoint on the RGB image, the depth data at time $\boldsymbol{t_n}$ are extrapolated from the 3D coordinates of the keypoint of the previous frame at time $\boldsymbol{t_{n-1}}$ by eq. (1).

$$\boldsymbol{P}(t_n) = \boldsymbol{P}(\boldsymbol{t_{n-1}}) + \frac{\boldsymbol{P}(\boldsymbol{t_{n-1}}) - \boldsymbol{P}(\boldsymbol{t_{n-2}})}{\boldsymbol{t_{n-1}} - \boldsymbol{t_{n-2}}}(\boldsymbol{t_n} - \boldsymbol{t_{n-1}}) \tag{1}$$

Finally, the sequence of the measured 3D skeleton of a person represents the motion of the person.

2.2 Learning 3D Human Motion

We use the recurrent neural network (RNN) described in[19] as our neural network model for both training and predicting 3D poses. The input consists of 3D pose data from the current and previous m frames ($(m+1)$ frames in total), while the network outputs 3D poses after T frames. Figure 2 illustrates the input and output data for the neural network model, where the data comprise the coordinates of the 3D keypoints in human motion. The number of nodes NN at the input and output layers is three times the number of keypoints, as each keypoint is represented by three data points (x, y, z) in a 3D space. In training, the 3D pose data $(\boldsymbol{P}(t_n), \ldots, \boldsymbol{P}(t_{n+m}))$ are sequentially used as input data frames to the model, while the 3D pose

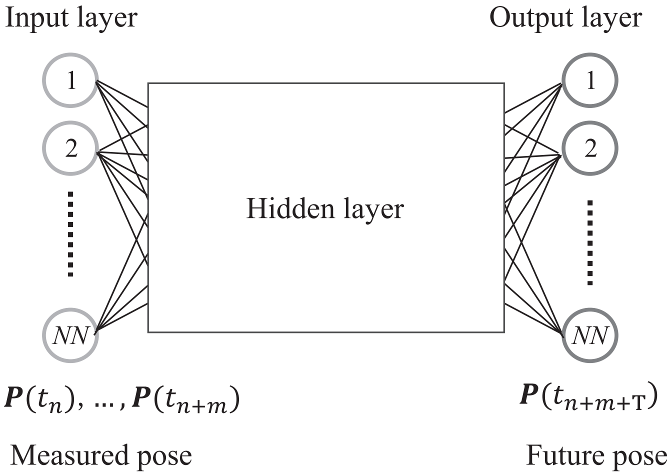

Input layer
Output layer
1
2
NN
Hidden layer
1
2
NN
$\boldsymbol{P}(t_n), \ldots, \boldsymbol{P}(t_{n+m})$
$\boldsymbol{P}(t_{n+m+\mathrm{T}})$
Measured pose
Future pose

Figure 2. Future pose predictor of RNN

Figure 2 is referred to in Section 3.2.

data after T data frames ($\boldsymbol{P}(t_{n+m+\mathrm{T}})$) serve as teaching data for the output layer. The mean squared error (MSE) is used as the error function to evaluate the error in the output layer.

The parameters for training the future pose predictor include the number of input data frames (m+1) in a sequence and the future frames T for prediction. The future frames T represent the number of frames after the last input data frame. The input data comprise successive (m+1) data frames, and the corresponding teaching data frame for prediction is a 3D pose after T frames from the last input data frame. The choice of the coordinate system for representing 3D human poses is also a parameter for training. Two coordinate systems, camera-centered and person-centered, are represented by C and P, respectively. Different future pose predictors are obtained by training with various combinations of parameters, including m, T, and the selected coordinate system.

2.3 Prediction of 3D Human Pose in Future Frame

For predicting the 3D human pose in future frames, the same RGB-D camera is used under identical environmental conditions as in the training stage. To predict the future 3D pose, a time sequence of a set of 3D keypoints of (m+1) data ($\boldsymbol{Pose}(t_i)$, ⋯, $\boldsymbol{Pose}(t_{i+m})$) measured in real time by the RGB-D camera is successively inputted into the neural network model's input layer. Then, the future pose predictor outputs the 3D pose ($\boldsymbol{Pose}(t_{i+m+T})$) represented by a set of 3D keypoints at a future frame $t_{i+m+\mathrm{T}}$ after T frames from the current time t_{i+m}. The successive future 3D pose is predicted by inputting the next data frame successively.

In[19], the predicted future 2D pose closely aligns with the measured 2D pose. However, based on examination of the prediction error in 2D and 3D poses, it is evident that both measurement and prediction errors are large in the depth direction, especially at keypoints near human boundaries such as the ears and tips of the feet.

3 Evaluation of 3D Pose Prediction for Different People

3.1 Learning 3D Motion of Different People and Prediction

The evaluation of the predicted 3D pose in future frames of a person is conducted in[19]. To clarify the effectiveness of training individual motions for predicting future poses, we train the future pose predictor for two different individuals, namely, A and B.

In the experiment, we use 17 keypoints of the skeleton, selected from the 25 keypoints detected by OpenPose. Keypoints corresponding to the ears or tips of the feet are excluded, as they do not accurately represent the dominant motion of a person, and their depth data are not reliable. The future pose predictors A and B share the same RNN structures, as illustrated in Fig. 2; however, the number of nodes NN at the input and output layers is 51 in the experiments. During the experiment, we use the Intel RealSense D455 as the RGB-D camera, positioned approximately 0.95 m above the floor.

For the collection and training of motion data, individuals A and B perform the same categorized predetermined actions. Each person performs 52 daily actions, such as ‘standing from a chair’ and ‘sitting on a chair’, in front of an RGB-D camera. Within the action data for each person, the 52 daily actions include 26 different types of predetermined actions that occur frequently in everyday activities, such as ‘standing from a chair’ and ‘sitting on a chair’. This set includes six instances of ‘standing from a chair’ and six instances of ‘sitting on a chair’. The number of data frames for an action is denoted by NF, with NF set to 60 in the experiments, and the iteration times N equal to 200 for training the future pose predictors of the model.

We obtain several motion predictors trained with various combinations of input and output parameters. A future pose predictor trained with different combination parameters is designated as (Coordinate system)_(number of input data frames). Here, ‘Coordinate system’ represents the camera-centered and person-centered coordinate systems denoted by C and P, respectively. In the person-centered coordinate system, the origin of the 3D keypoint coordinate system is set at the waist (mid-hip) of the person. For the ‘number of input data frames’ parameter, we chose $(m+1)$ to be equal to 3 and 5 for each coordinate system in the experiment. Furthermore, as a parameter indicating the number of future frames for prediction after the last input data frame, we set the number of future frames (T) to 3, 5, and 10. Combinations of the parameters include two coordinates C and P, two input data frames: 3 and 5, and three future frames: 3, 5, and 10. In total, we obtain 12 future pose predictors by training with these 12 different combinations of parameters.

In the experiments, the frame rate of the 3D pose prediction is approximately 10 fps. Thus, if the future frame T is set to 10, the motion predictor outputs the 3D pose at approximately 1.0 s into the future. To evaluate the prediction error of the 3D human pose for individuals in the future frame, we use two future pose predictors: A learns the motion of person A, and B learns the motion of person B. We evaluate the prediction errors when the predictor A predicts the motion of person A and person B, and when predictor B predicts the motion of person B and person A. Afterwards, we compare these prediction errors and discuss the effectiveness of training the motion of individual persons.

3.2 Evaluation of 3D Pose Prediction for Different People

For quantitative evaluation indices, we use mean per joint position error (MPJPE) and percentage of correct keypoints (PCK). MPJPE represents the mean distance between the predicted and measured 3D keypoints across the entire skeleton. In PCK, we determine the threshold based on the radius of the head, specifically using PCKh@0.5. PCKh@0.5 calculates the ratio of predicted keypoints within the radius of a sphere approximating the person's head. In our experiment, the radius is set to 0.110 m. To perform the evaluation, we compare and analyze these indices.

We evaluate the prediction errors for person A and B quantitatively; that is, we use two actions, 'standing from a chair' and 'sitting on a chair.' In each set of the training and prediction process for a future pose predictor, we select a pair of these actions (one 'standing' and one 'sitting') for testing, while using the remaining five pairs of actions and the remaining 50 action datasets for training. After completing this process for all six selected pairs of actions, we calculate the average MPJPE as the prediction error and the average PCKh@0.5 as the prediction precision for the motions of persons A and B.

Subsequently, we compare the prediction error and precision for the motions of two people using a future pose predictor trained by one of the two people. In the experiments, we call 'trainA' the future pose predictor trained by the motion of person A, and 'testA' represents the set of test actions for person A. We define 'trainB' and 'testB' in the same way. Next, we evaluate the prediction error for the six pairs of actions, including 'standing from a chair' and 'sitting on a chair'. For the evaluation, 'trainA' is applied not only to the test motion 'testA' of person A, but also to the test motion 'testB' of person B, and vice versa.

Moreover, we provide examples of predicted and measured skeletons for a sitting motion. In Fig. 3, a sequence illustrates the predicted pose of person A ('testA') by a future pose predictor trained by the motion of person A ('trainA'). Similarly, Fig. 4 presents a sequence

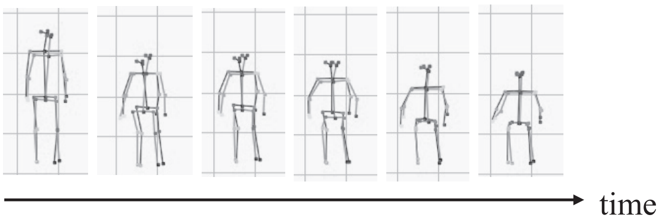


Figure 3. Example of predicted pose in sitting motion of person A ('testA' by 'trainA') (prediction: dark, measurement: bright)

Figure 3 is referred to in Section 3.2.

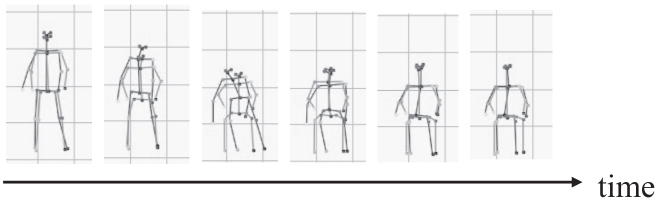


Figure 4. Example of predicted pose in sitting motion of person B ('testB' by 'trainA') (prediction: dark, measurement: bright)

Figure 4 is referred to in Section 3.2.

depicting the predicted pose of person B ('testB') by a future pose predictor trained by the motion of person A ('trainA'). The predicted skeleton is presented in dark color, while the measured skeleton is displayed in bright color.

These figures highlight that the prediction is more accurate in Fig. 3, where the persons are the same in training a future pose predictor and testing the motion. In Fig. 3, the predicted skeleton accurately matches the measured skeleton. However, in Fig. 4, differences between the predicted and measured poses are visible, particularly in the middle of the sitting motion. At the end of the sequence, the distance between the legs of the measured skeleton for person B is larger than that of the predicted legs. These prediction errors may be attributed to differences in individual sitting motions.

Next, Table 1 displays MPJPE as the prediction error for the average of standing and sitting actions, representing typical activities of daily living. Table 2 presents PCKh@0.5 as the prediction precision for the average of standing and sitting actions. The numerical values in the tables are the averages of the standing and sitting actions, where the values are averages of the six pairs of actions used for testing.

In Table 1, average MPJPE values under 100 mm are presented in bold. In Table 2,

Table 1. Average prediction error (MPJPE) for standing and sitting actions across six pairs of test motions

		Motion of person A (testA)			Motion of person B (testB)		
Future frames (T)		3	5	10	3	5	10
Future pose predictor (trainA)	C_3	**74.1**	**97.7**	133.7	131.3	179.0	251.2
	C_5	**71.9**	**87.9**	129.5	129.2	154.4	254.1
	P_3	**61.2**	**73.5**	**94.8**	**102.3**	120.1	146.9
	P_5	**60.1**	**67.8**	**83.0**	**102.1**	116.7	146.4
Future pose predictor (trainB)	C_3	132.3	162.0	256.5	**83.6**	114.4	177.2
	C_5	124.6	167.5	260.4	**85.4**	114.5	167.1
	P_3	**88.4**	**102.2**	136.1	**65.6**	**81.3**	111.9
	P_5	**86.8**	**103.7**	131.1	**66.4**	**77.5**	**97.9**

Table 1 is referred to in Section 3.2.

Table 2. Average prediction precision (PCKh @ 0.5) for standing and sitting actions across six pairs of test motions

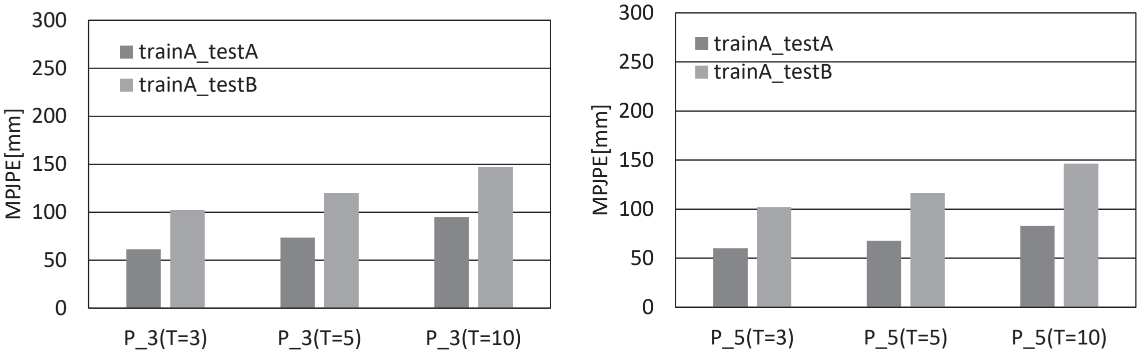
		Motion of person A (testA)			Motion of person B (testB)		
Future frames (T)		3	5	10	3	5	10
Future Pose predictor (trainA)	C_3	**82.4**	**69.3**	47.7	**53.0**	36.8	23.8
	C_5	**85.4**	**75.4**	**51.2**	**51.8**	45.1	19.4
	P_3	**89.4**	**82.3**	**70.0**	**65.8**	**59.5**	**51.1**
	P_5	**91.0**	**86.0**	**77.2**	**66.4**	**60.5**	**50.9**
Future pose predictor (trainB)	C_3	44.6	31.8	14.6	**77.8**	**62.8**	37.9
	C_5	49.4	32.4	16.7	**75.9**	**63.2**	44.2
	P_3	**72.2**	**64.4**	49.3	**87.1**	**78.7**	**59.9**
	P_5	**74.2**	**64.0**	**50.8**	**88.0**	**80.0**	**69.5**

Table 2 is referred to in Section 3.2.

average PCKh@0.5 values over 50% are presented in bold. Both tables demonstrate the prediction errors and precision for the combination of future pose predictors 'trainA' and 'trainB', and test motions 'testA' and 'testB'. Additionally, the tables illustrate that prediction is more accurate in the person-centered coordinate system than in the camera coordinate system. Furthermore, the experiments reveal a tendency of lower prediction error in near-future prediction.

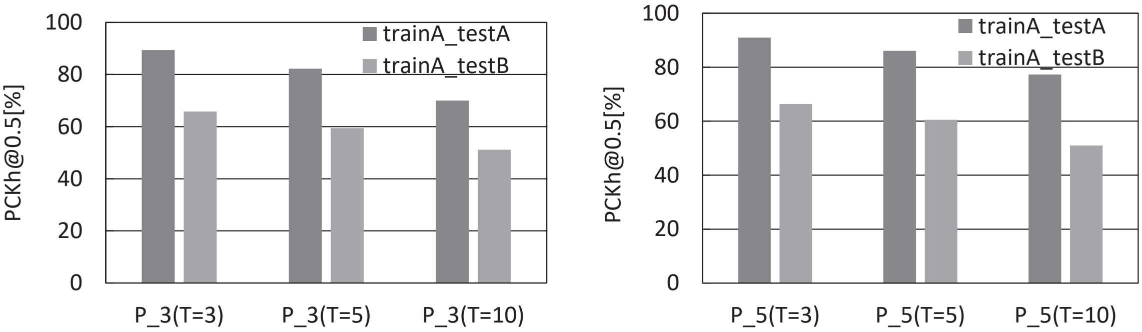
Figure 5 and 6 depict the prediction error (MPJPE) and precision (PCKh@0.5), respectively, for the motions of A and B predicted by the future pose predictor 'trainA' in the person-centered coordinate system. In these figures, 'trainA_testA' represents the case where the future pose predictor 'trainA' is applied to the motion of person A, and 'trainA_testB' represents the case where 'trainA' is applied to the motion of person B. These figures illustrate a consistent trend where prediction errors decrease and prediction precision values increase when the future pose predictor trained on the motion of a person is applied to the same person's motion.

Therefore, these results demonstrate the effectiveness of training a future pose predictor using the specific motion data of an individual to predict their future pose accurately.



(a) (b)

Figure 5. Average MPJPE of prediction for pose of person A and B by 'trainA'
Figure 5 is referred to in Section 3.2.



(a) (b)

Figure 6. Average PCKh@0.5 of prediction for pose of person A and B by 'trainA'
Figure 6 is referred to in Section 3.2.

4 Conclusion

We describe a method for the real-time prediction of 3D human pose using an RGB-D for personal applications. Our approach demonstrates the distinct differences in motion among individuals and the effectiveness of using personal motion data for learning and predicting individualized motions. Using an RGB-D camera, we captured 3D poses and trained an RNN to predict future 3D poses based on a sequence of input data frames in real time. Next, we evaluated the prediction error for two representative daily actions, namely, ‘sitting on a chair’ and ‘standing from a chair’. Despite the limited size of the training dataset for each person, the predicted future pose closely approximated the measured pose. The results indicate the effectiveness of training the predictor using an individual’s motion data for accurately predicting their motion.

Future work will involve evaluating differences in individuals’ motions across a broader range of everyday activities. Creating datasets that include diverse motions for individuals through long-term observations is also on the agenda. Additionally, employing larger action datasets from real-work everyday environments during draining is expected to enhance the precision of future pose predictions. This advancement holds promise for the application of motion predictions to service robots in real-world scenarios, enabling greater personalization in their support for individual users.

Acknowledgement

This work was financially supported in part by the Kansai University Fund for the Promotion and Enhancement of Education and Research, 2020. “Real-world Service Innovation through AI robot challenges.”

References

1) Pishchulin, L., Insafutdinov, E., Tang, S., Andres, B., Andriluka, M., Gehler, P., Schie-le, B.: DeepCut: Joint subset partition and labeling for multi person pose estimation. In: CVPR2016, pp. 4929-4937 (2016)

2) Cao, Z., Simon, T. Wei, S., Sheikh, Y.: Realtime multi-person 2D pose estimation using part affinity fields. In: CVPR2017, pp. 1302-1310 (2017)

3) Papandreou, G., Zhu, T., Chen, L., Gidaris, S., Tompson, J., Murphy, K.: PersonLab: Person pose estimation and instance segmentation with a bottom-up, part-based, geo-metric embedding model. In: Ferrari,V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV2018. LNCS, vol. 11218, pp. 282-299. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01264-9_17

4) T. L. Munea, Y. Z. Jembre, H. T. Weldegebriel, L. Chen, C. Huang and C. Yang, “The progress of human pose estimation: A survey and taxonomy of models applied in 2d human pose estimation,” in IEEE Access, vol. 8, pp. 133330-133348, (2020)

5) Chen, C.-H., Ramanan, D.: 3D human pose estimation = 2D pose estimation + matching. In: CVPR2017, pp. 5759-5767 (2017)

6) Pavlakos, G., Zhou, X., Derpanis, K.G., Daniilidis, K.: Coarse-to-fine volumetric prediction for single-image 3D human pose. In: CVPR2017, pp. 1263-1272 (2017)

7) Zhou, X., Huang, Q., Sun, X., Xue, X., Wei, Y.: Towards 3D human pose estimation in the wild: A weakly-supervised approach. In: ICCV2017, pp. 398-407 (2017)

8) Novotny, D., Ravi, N., Graham, B., Neverova, N., Vedaldi, A.: C3DPO: Canonical 3D pose networks for non-rigid structure from motion. In: ICCV2019, pp. 7687-7696 (2019)

9) Cheng, Y., Yang, B.,Wang, B., Yan, W., Tan, R.T.: Occlusion-aware networks for 3D human pose estimation in video. In: ICCV2019, pp. 723-732 (2019)

10) Moon, G., Chang, J.Y., Lee, K.M.: Camera distance-aware top-down approach for 3D multi-person pose estimation from a single RGB image. In: ICCV2019, pp. 10132-10141 (2019)

11) Sarafianos N, Boteanu B, Ionescu B, Kakadiaris IA: 3d human pose estimation: A review of the literature and analysis of covariates. Computer Vision and Image Understanding 152:1-20 (2016)

12) Chen Y, Tian Y, He M: Monocular human pose estimation: A survey of deep learning-based methods. Computer Vision and Image Understanding 192:102897 (2020)

13) Martinez, J., Hossain, R., Romero, J., Little, J.J.:A simple yet effective baseline for 3D human pose estimation. In: ICCV2017, pp. 2659-2668 (2017)

14) Chiu, H., Adeli, E., Wang, B., Huang, D., Niebles, J.C.: Action-agnostic human pose forecasting. In: 2019 IEEE Winter Conference on Applications of Computer Vision, pp. 1423-1432 (2019)

15) Wu, J., Xue, T., Lim, J.J., Tian, Y., Tenenbaum, J.B., Torralba, A., Freeman, W.T.: Single image 3D interpreter network. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9910, pp. 365-382. Springer, Cham (2016) https://doi.org/10.1007/978-3-319-46466-4_22

16) Newell, A., Yang, K., Deng, J.: Stacked hourglass networks for human pose estimation. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9912, pp. 483-499. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46484-8_29

17) Chao, Y.-W., Yang, J., Price, B., Cohen, S., Deng, J.: Forecasting human dynamics from static images. In: CVPR2017, pp. 3643-3651 (2017)

18) Horiuchi, Y., Makino, Y., Shinoda, H.: Computational foresight: Forecasting human body motion in real-time for reducing delays in interactive system. In: Proceedings of 2017 ACM International Conference on Interactive Surfaces and Spaces, pp. 312-317 (2017)

19) Mae, Y., Nagata, A., Tsunoda, K, Takahashi, T., Suzuki, M., Arai, Y., Aoyagi, S.: Real-time prediction of future 3D pose of person using RGB-D camera for personalized services. In: Huang DS., Jo KH., Li J., Gribova V., Bevilacqua V. (eds) Intelligent Computing Theories and Application. ICIC 2021. Lecture Notes in Computer Science, vol 12836. Springer, Cham. https://doi.org/10.1007/978-3-030-84522-3_69