


Received April 26, 2020, accepted May 13, 2020, date of publication May 25, 2020, date of current version June 8, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2997341

Using Sensors and Deep Learning to Enable On-Demand Balance Evaluation for Effective Physical Therapy

WENCHUAN WEI¹ , (Member, IEEE), CARTER MCELROY²,
AND SUJIT DEY¹, (Fellow, IEEE)

¹Electrical and Computer Engineering Department, University of California San Diego, La Jolla, CA 92093, USA

²Department of Rehabilitation Services, University of California San Diego, La Jolla, CA 92093, USA

Corresponding author: Wenchuan Wei (w8wei@eng.ucsd.edu)

This work was supported by the National Science Foundation under Grant IIS-1522125.

ABSTRACT In traditional physical therapy, balance evaluation is performed by the physical therapist (PT) intermittently during clinic visits, which is subjective, inconvenient, and time-consuming. In this paper, we use sensors and deep learning to propose an automated balance evaluation system for home and clinical use. First, we propose a deep learning-based model to estimate the subject's Center of Mass (CoM) position using a depth camera, which outperforms other CoM estimation methods with high accuracy and ease of use. Then we propose a balance evaluation system to evaluate the subject's dynamic balance in a Gait Initiation (GI) task. The subject's CoM position is estimated by the proposed CoM estimation model and the Center of Pressure (CoP) position is measured by a Wii balance board. The CoP-CoM trajectory during the GI task is used to assess and quantify the patient's dynamic balance control. Using data collected from both healthy subjects and patients with Parkinson's Disease, the proposed balance evaluation model is able to quantify the subject's balance level which is consistent with the human PT's assessments in traditional balance evaluation tests. The proposed balance evaluation system can be used as a portable and low-cost tool for on-demand balance evaluation.


INDEX TERMS Balance evaluation, biomedical informatics, computer vision, deep learning, machine learning, patient rehabilitation, physical therapy, gait initiation.

I. INTRODUCTION

In physical therapy, the patient's ability to balance is an important indicator for the physical therapist (PT) to select the proper training programs, evaluate the progress of the patient, predict fall risk [1], etc. Traditionally, balance evaluation is performed by the PT at the initial evaluation and intermittently during clinic visits. However, the patient's balance may change over time and also be influenced by medication, sleep quality, etc. Therefore, it is important to have more frequent and preferably on-demand balance evaluation to monitor the patient's condition. Moreover, traditional balance evaluation tests like the Berg Balance Scale (BBS) [2] and the mini Balance Evaluation Systems Test (mini-BESTest) [3] are time-consuming and require the PT's subjective assessments, therefore they may be limited for

clinical use. To address the problems of traditional balance evaluation, Mishra *et al.* have proposed to use a camera system to evaluate the static balance (i.e., the ability to stay stationary in some postures) using static body sway in single-leg stance [4]. For dynamic balance (i.e., the ability to maintain balance in motion or recover from imbalanced conditions), Kennedy *et al.* have proposed the WeHab system to measure the patient's balance in dynamic tasks (e.g., sit-to-stand and weight-shifting) but do not achieve good results [5]. In this paper, we focus on the dynamic balance evaluation for patients with Parkinson's disease (PD) as dynamic balance is more important to improve agility and avoid falls. We propose an automated balance evaluation system using multiple sensors and deep learning to provide accurate, convenient, and on-demand balance evaluation for home and clinical use.

In balance evaluation, an important indicator is the Center of Mass (CoM) position of the human body. For the 3D position of the human's CoM, the horizontal CoM (i.e., the

The associate editor coordinating the review of this manuscript and approving it for publication was Wenbing Zhao .

projection of CoM on the ground) is of greater importance [4], [5]. Since the CoM position of the human body cannot be directly measured, researchers have proposed to measure the Center of Pressure (CoP) of the ground reaction force in static/balanced postures to represent the horizontal CoM position [10]–[12], [20]. In a static/balanced posture (e.g., quiet standing), the only forces acting on the human body are the gravity (which acts on the CoM) and the ground reaction force. According to Newton's second law, the gravity is equivalent to the ground reaction force in both magnitude and position (i.e., $\text{CoP} = \text{horizontal CoM}$) since the acceleration of the human body is zero in static/balanced postures. The traditional way to measure the CoP position is using the laboratory-grade force plate. However, the force plate is primarily limited to laboratory use due to its high cost and complicated setup procedure. The Wii Balance Board (WBB) is a device designed by Nintendo for balance-related games and can calculate the CoP position of the human body. The CoP measurement error of the WBB has been proved to be within 5 mm [6]. Because of its low cost, portability, and high accuracy in CoP measurement, the WBB has been increasingly used as a replacement of the force plate in many studies [6]–[8].

However, the CoP position measured by the force plate or the WBB is equivalent to the horizontal CoM position only when the user is in a static/balanced posture. Moreover, the force plate or the WBB needs to be placed on a horizontal and firm plane to measure the CoP position accurately. In balance evaluation, we often need to test the subject's dynamic balance or the subject's static balance on different surface types (e.g., the incline ramp, or the foam). To solve this problem, researchers have proposed to use pose and body parameters (e.g., body shape and density) to estimate the CoM position. In previous studies, the body shape of the subject is either modeled as geometrical segments [9], [10] or estimated from an identification/calibration process [11], [20]. To achieve identification-free CoM estimation, Kaichi *et al.* have proposed a voxel reconstruction approach to reconstruct the subject's 3D body using multiple cameras and estimate the CoM position by assigning weights to the body parts [12]. However, they need to carefully calibrate five cameras for the 3D body reconstruction, which makes it not suitable for home and clinical use.

In recent years, vision-based models have been increasingly used to learn and predict human-related activities, for example, facial expression recognition [13], fall prediction [14], etc. Inspired by these techniques, we propose to use deep learning to learn the body parameters of the subject and estimate the horizontal CoM position. We have selected the depth camera instead of an RGB camera because the depth map provides more information about the subject's body in the depth direction, which is essential in CoM estimation. Besides, depth cameras work better in low light conditions and are color and texture invariant [15]. Figure 1 shows the proposed CoM estimation model. Motivated by the use of Convolutional Neural Network (CNN) in pose estimation

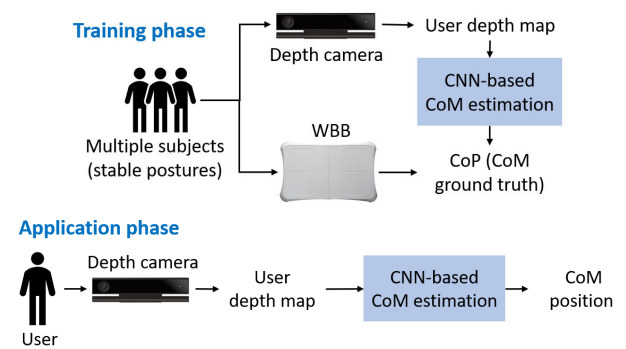


FIGURE 1. The proposed CoM estimation model: training and application phase.

problems [35], [36], we propose to use CNN in our CoM estimation model as estimating the human's CoM position is similar to estimating the joint positions (i.e., pose estimation). In the training phase, a CNN-based model is trained using data collected from multiple subjects in various static postures. We use a depth camera to capture the depth images and a WBB to measure the ground-truth CoP position. In the application phase, only the depth camera is needed to estimate the subject-specific CoM position. The depth camera is anyway necessary in most automated training systems for its ability in skeleton tracking and motion capture [16], [17]. By using the proposed CoM estimation model, the subject's CoM position can also be tracked without any extra device.

Note that the CoM estimation model is trained from data collected in static postures and it will be used for dynamic postures in the balance evaluation system. Despite the fact that there is no direct way to validate its accuracy on dynamic postures (as the ground truth of CoM position cannot be measured), we will demonstrate that the balance evaluation model built upon the CoM estimation model is able to provide accurate balance assessments that are consistent with the PT score. Therefore, it is reasonable to conclude that the proposed CoM estimation model can provide accurate CoM estimation for both static and dynamic postures. By using a single depth camera that does not need complicated setup or subject identification, the proposed CoM estimation model can be used as a portable and low-cost tool for subject-specific CoM measurements.

Based on the CoM estimation model, we further propose the balance evaluation system using multiple sensors. The tested task is Gait Initiation (GI), which refers to the transient period between the quiet standing posture and steady state walking. Patients with impaired balance have difficulty in performing the correct body weight shift in GI [30]. Hass *et al.* have proposed that the CoP-CoM distance during GI is an important indicator of dynamic balance control [18]. Inspired by their research, we propose to develop an automated balance evaluation system to provide quantitative balance evaluation using the GI task and mimic the human PT's assessments during traditional balance tests. The proposed system is shown in Figure 2. The depth camera and the WBB measures the subject's CoM and CoP positions during

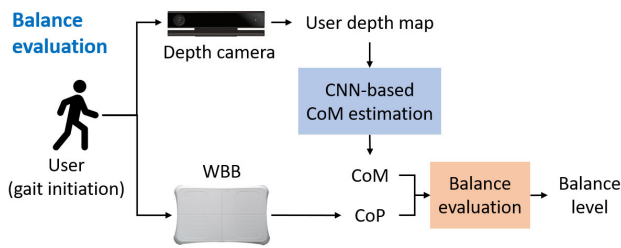


FIGURE 2. The proposed balance evaluation system.

the GI task respectively. The patient's balance level will be calculated based on the CoP and CoM trajectory. To the best of our knowledge, our proposed system is the first to provide automated and quantitative evaluation on the subject's dynamic balance, which can mimic the human PT's manual assessments in the mini-BESTest. While we focus on patients with Parkinson's disease (PD) in this paper, the proposed balance evaluation system can be used in the physical therapy for any disease/condition where balance evaluation is critical (e.g., orthopedic disease and stroke).

A preliminary version of this work has been reported in [19], which introduced a CoM estimation model. However, the model proposed in [19] did not show high accuracy. In this paper, we develop an enhanced CoM estimation model by using colored skeleton images instead of joint heatmaps as inputs to the model, and proposing a novel coarse-to-fine approach to improve the accuracy. The enhanced model reduces the estimation error by about 10%, compared with the preliminary model in [19]. Moreover, our preliminary work [19] proposed only the CoM estimation model, whereas this paper uses the enhanced CoM estimation model to propose an automated balance evaluation system, which for the first time enables quantitative, accurate, and on-demand dynamic balance evaluation for home and clinical use. Compared with traditional balance evaluation (e.g., the mini-BESTest conducted by a PT, or tests using laboratory-grade devices), the proposed balance evaluation system can be used at home or away (e.g. at hotels while traveling), or in clinics without PTs or high-end devices (e.g., retail-based clinics and mobile clinics). The patients can use the proposed system as a portable and low-cost tool to measure their balance on an on-demand basis, which enables closer monitoring of their health condition and progress in physical therapy training. The proposed balance evaluation system has the potential of significantly reducing PT visit requirements and reducing cost for both the patients and care providers. The main contributions of this paper can be summarized as follows:

- We have proposed a CNN-based CoM estimation model to estimate a subject's CoM position from a single depth image. Compared with other CoM estimation methods, the proposed approach does not need any subject identification process and can estimate the subject-specific CoM position with high accuracy, which is convenient for home and clinical use.
- We have proposed to use the colored skeleton map instead of the joint heatmaps (proposed in [19]) as the

input of the CNN model. The colored skeleton image reduces the training and inference times of the CNN model significantly by reducing the input dimension and the number of parameters of the model.

- To solve the trade-off problem in the selection of the discretization interval (DI) when discretizing the CoM coordinates, we have proposed the novel coarse-to-fine approach to improve the accuracy of the CoM estimation model.
- We have proposed a balance evaluation system using inexpensive and portable sensors (i.e., a depth camera and a WBB) to measure the subject's balance level during a simple GI task. To the best of our knowledge, our proposed system is the first to provide automated and quantitative evaluation on the subject's dynamic balance, which can mimic the human PT's manual assessments in the mini-BESTest.

The rest of the paper is organized as follows: Section II introduces the related work on CoM estimation and balance evaluation in more details. In Section III, we introduce the methods used in the proposed models, including the CoM estimation model in Section III-B and the balance evaluation system in Section III-C. Section IV describes the experimental results. Section V concludes the paper and discusses future work.

II. RELATED WORK

While we have briefly discussed the related work on CoM estimation and balance evaluation in the previous section, we next explain the most relevant techniques in more details, pointing out their disadvantages and the need and differentiation of our proposed technique.

A. RELATED WORK ON CoM ESTIMATION

1) CoM ESTIMATION USING IMU SENSORS [33], [34]

Some studies used Inertial Measurement Unit (IMU) sensors to estimate the CoM position. Esser *et al.* proposed to estimate the subject's vertical CoM movements from the acceleration data collected by the IMU sensor by [33]. However, the wearable IMU sensors are not convenient for patients with impaired mobility.

2) WINTER'S METHOD

Winter proposed a kinematic method to estimate the CoM position of the human body [9]. He modeled the human body as 16 segments and used a motion capture system to track the position of each segment. The CoM position of the whole body was calculated as the weighted sum of the CoM position of each segment. The weight of each segment was taken from previous anthropometric studies. However, this method cannot provide subject-specific CoM estimation as the weight of each segment may differ in subjects of different age, sex, and fitness level, etc.

3) THE OPTIMIZATION-BASED METHOD

Chen *et al.* proposed to use an optimization-based model to estimate the body parameters of the subject [10]. They

modeled the human body as some geometric shapes and measured the size of each segment manually. A force plate was used to measure the CoP position as the ground truth of the horizontal CoM position. However, modeling the body segments as geometrical shapes (e.g., modeling the neck as a frustum) is not accurate and the manual measurement of the body size is inconvenient.

4) THE STATICALLY EQUIVALENT SERIAL CHAIN (SESC) MODEL

The SESC model translates the human's mass distribution to the geometry of a linked chain [11]. An identification phase was used to obtain the subject-specific SESC parameters. In the identification phase, each subject performed 14 static postures. Later, Gonzalez *et al.* proposed that using more postures in the identification phase and assuming the bilateral symmetry of the human body can reduce the estimation error of the SESC method [20]. They also showed that using low-cost sensors Kinect and WBB can achieve comparable results to those obtained using high-end equipment. However, the subject identification phase still needs to be conducted each time when a new subject comes or the mass distribution of an existing subject has changed, which limits its application.

5) THE VOXEL RECONSTRUCTION METHOD

Kaichi *et al.* proposed to reconstruct the 3D human body and then estimate the CoM position [12]. They used five cameras to capture multiple views of the human body and a 3D reconstruction approach to reconstruct the body. The human body was segmented into nine parts and the CoM position of the whole body was estimated as the weighted sum of the position of each part. The weights were taken from previous anthropometric studies. As mentioned in Section I, the main challenges in the subject-specific CoM estimation problem include the difference in body size and density. By reconstructing the 3D body, the voxel reconstruction approach solves the problem of difference in body size but still fails to consider the difference in body density since it uses the density information from previous studies. Moreover, the five cameras need to be carefully calibrated. In comparison, our proposed model uses a single depth camera and does not need any complicated calibration or subject identification process, which is more convenient for home and clinical use.

B. RELATED WORK ON BALANCE EVALUATION

The balance control of the human body includes static balance and dynamic balance. Static balance refers to the ability to stay stationary in some postures (e.g., single-leg stance), while dynamic balance refers to the ability to maintain balance in motion or recover from imbalanced conditions. For static balance evaluation, the body sway during single or two-legged stance is used. The body sway is presented by the moving range of the CoM positions, which can be measured by a force plate or a WBB (as CoP = CoM in static conditions) [21], or estimated using the above CoM estimation methods [4]. Subjects with better static balance would

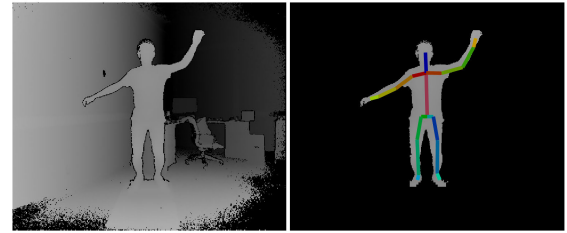


FIGURE 3. Left: the original depth map captured by the depth camera of Kinect. Right: the user depth map and the colored skeleton image overlay.

have smaller body sway. For dynamic balance, Hsu *et al.* proposed to use an inertial-sensor-based wearable device to analyze gait information and balance ability for patients with Alzheimer's disease [32]. However, wearable sensors attached on the body may cause extra burden to the users, especially for patients with impaired mobility. Therefore, we decide to use non-wearable sensors (e.g., cameras and balance boards) in the proposed balance evaluation system for patients with PD. Hass *et al.* proposed that the CoP-CoM distance during the GI task might represent the dynamic balance control of patients with PD and shown that the peak magnitude of the CoP-CoM distance was smaller in more balance-impaired patients than in healthy subjects. However, the CoM measurements in their work were based on the skeleton-based approach [9] and were not accurate. Moreover, they provided only qualitative results by showing the difference in CoP-CoM distance between patients with PD and healthy subjects. In comparison, our proposed balance evaluation model is able to provide quantitative balance level, which is consistent with the human PT's manual assessments in standardized balance tests. The quantitative balance level can be used to select the proper training programs, evaluate the patient's progress, and predict the fall risk.

III. METHODS

A. DEVICES: KINECT AND WII BALANCE BOARD

The Kinect sensor can capture the human pose using an RGB camera and a depth camera [22]. Each pixel in the depth map represents the distance of the pixel from the sensor. Based on the original depth map, the user depth map (by removing the background) and the user skeleton can be obtained [15] (see Figure 3).

The Wii balance board (WBB) consists of four pressure sensors located at the four corners of the board. When a user stands on the board, the four pressure sensors measure the vertical force and the CoP can be calculated. Compared with our preliminary work in [19], we have extended the range of CoP measurements by using two WBBs side by side to enable more postures. Figure 4 shows the two WBBs and the coordinate system. In this paper, the x - and y -axis are defined as the length and width direction of the WBB, and the z -axis is the upright direction. Based on torque equilibrium, the CoP position can be calculated as

$$x = \frac{L}{2} \times \frac{(P_{12} + P_{14} + P_{22} + P_{24}) - (P_{11} + P_{13} + P_{21} + P_{23})}{P_{11} + P_{12} + P_{13} + P_{14} + P_{21} + P_{22} + P_{23} + P_{24}}, \quad (1)$$

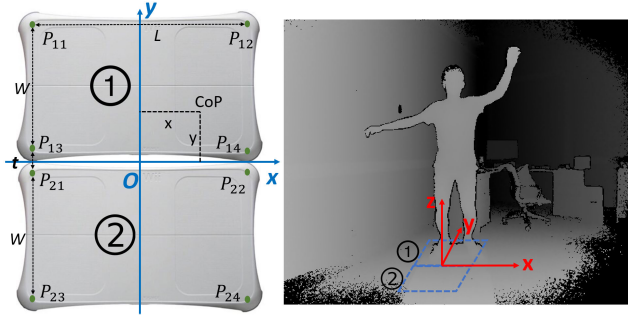


FIGURE 4. Two WBBs and the 3D coordinate system.

$$y = \frac{(t+W)(P_{11}+P_{12}-P_{23}-P_{24})+t(P_{13}+P_{14}-P_{21}-P_{22})}{P_{11}+P_{12}+P_{13}+P_{14}+P_{21}+P_{22}+P_{23}+P_{24}}, \quad (2)$$

where L and W are the length and width of the board, t is the size of the gap between the two boards, and P_{ij} is the force measured by the j -th pressure sensor of the i -th board. Several studies have found that the CoP measurement error of the WBB is smaller than 5 mm, compared with the laboratory-grade force plate [6], [23]. Besides, the WBB is inexpensive and portable, which makes it a good tool for home and clinical use. Therefore, we have selected the WBB to measure the CoP positions in this paper.

B. THE PROPOSED CoM ESTIMATION MODEL

1) INPUT AND OUTPUT OF THE MODEL

For the CoM estimation model, the input is the full depth map and the output is the horizontal CoM position of the user. To help the model distinguish between different body parts (as different parts may have different densities), we have proposed in our preliminary work [19] to use the joint heatmaps to provide information about the joint positions. However, the joint heatmaps are high-dimensional and introduce too many parameters in the CNN model. As the heatmap of each joint has the same size as the input depth image (512×424), the heatmaps of all 25 joints have 25 channels ($512 \times 424 \times 25$). To reduce the number of parameters in the CNN model, we further propose to use the colored skeleton image instead of the joint heatmaps to provide information about the different body parts of the subject. The colored skeleton image is created by connecting the adjacent joints of the body and using a specific color for each body segment. For example, the right shank connecting the right knee joint and the right ankle joint is rendered in light blue (RGB = [0, 102, 153]). Figure 3 shows an example of the colored skeleton. The colored skeleton image also has the same size as the depth image (512×424) but has only 3 channels, compared with the 25 channels of the joint heatmaps proposed in [19]. Therefore, the colored skeleton image can reduce the training and inference times of the CNN model by reducing the input dimension and the number of parameters of the model. Each body segment is rendered in a different color so the network

can differentiate between different body parts. The user depth map and the colored skeleton image are concatenated as the input of the CNN model.

The output of the model is the horizontal CoM position of the user. As shown in (1) and (2), the horizontal CoM positions measured by the WBB are continuous values (x, y), therefore the CoM estimation is a regression problem. However, it has been proved that the direct regression of coordinates from images is a highly non-linear problem and learning the mapping is a challenging task [24]. To solve this problem, we propose to discretize the continuous coordinates into discrete classes. For each data sample, the CNN model will predict the most likely discretized class k_x and k_y ($k_x, k_y = 0, 1, 2, \dots$) and the continuous CoM coordinate will be estimated as the center of the discretized class as

$$CoM_x = (k_x + 0.5) \times I_x, CoM_y = (k_y + 0.5) \times I_y \quad (3)$$

where I_x and I_y are the length of the discretization interval (DI) in the x - and y - direction. More details about the selection of DI will be discussed in Section III-B4. By discretizing the continuous CoM coordinates, we cast the highly non-linear problem of direct CoM coordinate regression to a more manageable form of classification in a discretized space.

2) DATA AUGMENTATION

Data augmentation is an important step in deep learning to increase the amount and diversity of the training data and reduce overfitting. Traditional data augmentation approaches include rotating, flipping, translating the image, and/or adding noise to the image. In image classification, these operations are useful as they do not change the image categories. However, they cannot be directly applied to our dataset as the CoM position of the user may be different. To solve this problem, we propose to apply different data augmentation approaches to the x - and y -component of the CoM position separately.

For the x -component of the CoM position, two data augmentation approaches are applied to the user depth map: (1) Adding a random depth value to the user body area, which is identical to shifting the user body in the depth direction. (2) Shifting the user body randomly in the z -direction. Both operations will not change the x -value of the CoM position. For the y -component of the CoM position, two data augmentation approaches are applied to the user depth map: (1) Shifting the user body in the x -direction randomly. (2) Shifting the user body in the z -direction randomly. Both operations will not change the y -value of the CoM position. Note that the colored skeleton images also need to be processed in the same way as the user body (i.e., adding the same depth value and shifting the same amount).

3) CNN-BASED NETWORK ARCHITECTURE

In computer vision problems, CNN [25] is widely used for its advantages in feature extraction, parameter sharing, etc. We propose a CNN-based model for the CoM estimation

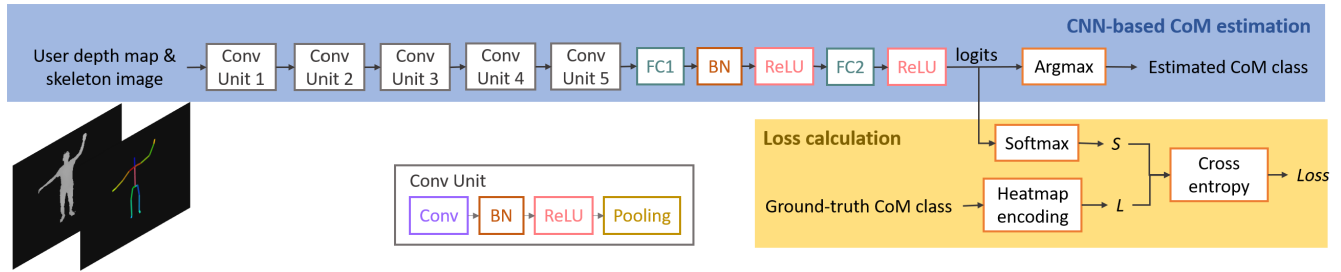


FIGURE 5. The proposed CNN model for CoM estimation.

problem (see Figure 5). In each convolutional unit, we use a Convolutional (Conv) layer [26] to extract features from the original image or the output of the previous layer, a Batch Normalization (BN) layer [27] to stabilize the inputs to the following nonlinear activation function, a Rectified Linear Unit (ReLU) layer to add non-linear transformation, and a max Pooling layer to reduce the size of each feature map. We use five Conv units to extract features from the depth images. The number of layers is selected empirically and details about our implementation are shown in Section IV-B. After the five Conv units, we use two Fully Connected (FC) layers to output the probability of each discrete CoM class from the results of previous Conv units, and an Argmax layer to select the final output with the highest probability. As described in Section III-B1, the continuous CoM positions have been discretized into some classes, so the CNN model will do a classification to decide the correct class of the CoM coordinates. We define the loss function as the cross-entropy between the ground-truth class of the CoM and the predicted CoM class as follows.

$$Loss = - \sum_{i=1}^N L_i \log(S_i), \quad (4)$$

where L_i is the encoding for class i in the ground-truth CoM and S_i is the softmax output of class i in the estimated CoM. In most image classification problems, the traditional encoding method for the ground-truth label is one-hot encoding as follows.

$$L_i = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases} \quad (5)$$

where k is the ground-truth class. In this way, the ground-truth class k is encoded as 1 and all the other classes are encoded as 0. Figure 6 shows an example. One-hot encoding is used in image classification problems because the label for an image is a categorical feature and all the incorrect classes ($i \neq k$) should be considered equally. However, the ground-truth class of CoM position is discretized from the continuous value, so the incorrect classes should be penalized differently according to their distance to the ground-truth class. Thus, we propose to use Gaussian-distributed heatmap instead of

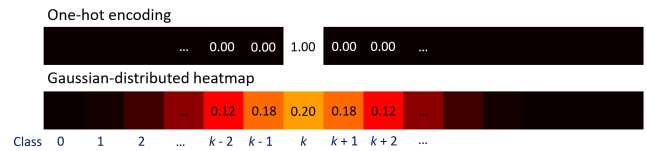


FIGURE 6. CoM ground-truth class encoding (k is the true class): one-hot encoding and Gaussian-distributed heatmap.

one-hot encoding to encode the ground-truth CoM as

$$L_i = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(i-k)^2}{2\sigma^2}}, \quad (6)$$

where σ is the standard deviation of the Gaussian distribution. An example of the Gaussian heatmap is also shown in Figure 6. The ground-truth class k has the highest probability 0.20 and the other classes are encoded according to their distance to the ground-truth class k . The CoM heatmap represents the confidence of each class as the ground truth. By using the Gaussian heatmap, the CNN model can be trained to move its output towards the ground-truth class during the learning process.

4) A COARSE-TO-FINE APPROACH TO INCREASE THE ACCURACY

As discussed in Section III-B1, the continuous CoM coordinates are discretized into some classes in the CNN model. However, there are some trade-offs in the selection of the discretization interval (DI) when discretizing the CoM coordinates. Smaller DI leads to larger number of discretized classes and therefore more challenges in the classification problem due to some outliers. Figure 7 shows an example. The numbers in each block represent the output probability of each class. The outlier class has a probability 0.16, which is higher than the correct class (probability = 0.15). For larger DI, there are smaller number of classes, which leads to higher accuracy in the classification problem. However, the final CoM estimation error may still be high as the true CoM position within the class may be far from the center of the interval that is estimated as the output CoP position.

To solve the above problems, we propose a coarse-to-fine approach to avoid outliers and improve the accuracy in CoM estimation. First, we train several CNN models with different

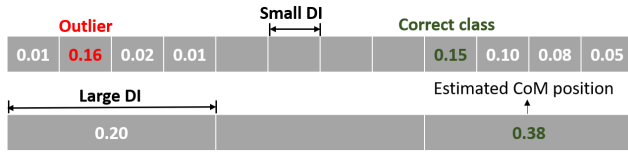


FIGURE 7. Trade-off on the selection of discretization interval (DI).

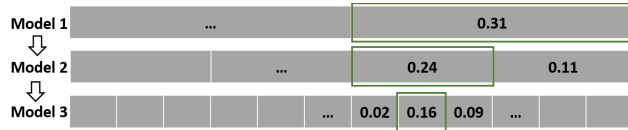


FIGURE 8. An example of the proposed coarse-to-fine approach. The green box represents the selected class in each model.

DIs in descending order ($DI_1 > DI_2 > \dots$) and DI_k should be a multiple of DI_{k-1} (i.e., $DI_k = m_k \times DI_{k-1}$ where m_k is an integer). As larger DI ensures higher accuracy in the classification problem, we first use the model with the largest DI to decide the coarse range of the CoM position. Then, instead of directly using the center of the interval as the output, we use the model with smaller DI to obtain finer estimation of the CoM position. Figure 8 shows an example of three models ($DI_1 = 2DI_2 = 6DI_3$). We start with Model 1 and select the class with the highest probability (shown in green box). Then we use Model 2 and select between the two sub-classes that lie in the selected range resulting from Model 1. Similarly, we use Model 3 and select between the three sub-sub-classes that lie in the selected range resulting from Model 2. In this way, the outliers that may exist in the fine model (with small DI) are excluded in the coarse model (with large DI) and the precision of CoM estimation is improved in each step as the DI goes smaller. For the last model (with the smallest DI), we will output the final CoP position as the center of the selected small interval. Although the inference time will increase by using multiple models in the proposed coarse-to-fine approach, the inference time of each model is negligible (about 13 ms, see Table 2) by using the proposed colored skeleton image (proposed in Section III-B1 and validated in Section IV-C). Therefore, the total inference time on multiple models is also very small (<40 ms, see Table 2) by using the proposed coarse-to-fine approach.

C. THE PROPOSED BALANCE EVALUATION SYSTEM

Based on the CoM estimation model, we further propose a balance evaluation system to provide quantitative balance evaluation using the GI task. The subject's depth images and CoP positions are captured by the Kinect camera and the WBB (see Section III-A). The subject's CoM positions are estimated from the depth images using the proposed CoM estimation model. As GI is a dynamic posture, the subject's CoP position is not equivalent to the CoM position. As proposed in [18], the maximum distance between the subject's CoP and CoM position during GI is correlated with the subject's dynamic balance control. Therefore, we calculate the CoP-CoM distance during the GI task. An example of the

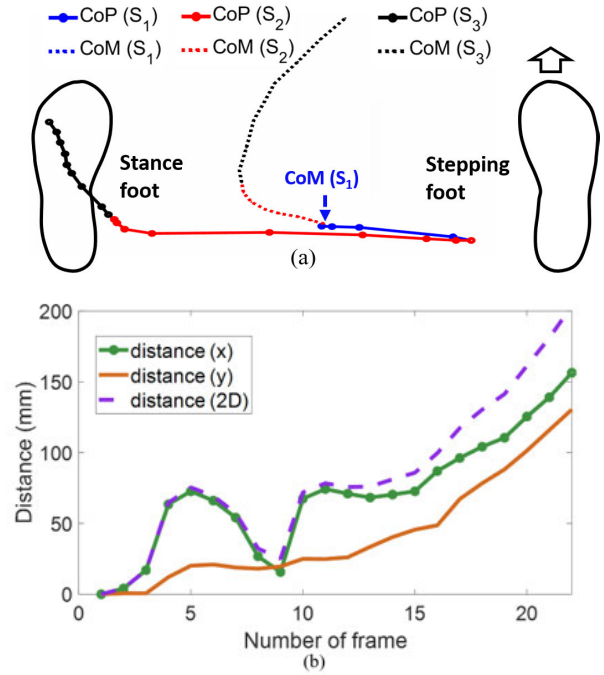


FIGURE 9. (a) The CoP-CoM trajectory during the GI task in three states. S_1 : CoP shifts towards the stepping foot and CoM remains at the original position; S_2 : CoP shifts back towards the standing limb; S_3 : both CoP and CoM move forward. (b) The CoP-CoM distance vs. frame number during the GI task.

CoP-CoM trajectory and the CoP-CoM distance vs. time in the x/y direction and the 2D distance (i.e., distance in the xy plane) during GI is shown in Figure 9. The right foot is the stepping foot. The subject's motion during GI can be divided into three states $S_1 \sim S_3$. In S_1 , the CoP of the subject shifts towards the stepping foot and the CoM remains at the original position, therefore the CoP-CoM distance increases. In S_2 , the subject's CoP shifts back towards the standing limb, as the stepping limb advances. During this time, the CoP-CoM distance first decreases and then increases. In S_3 , the subject's CoP and CoM both move forward and the CoP-CoM distance continues to increase. From Figure 9 we can see that the maximum CoP-CoM distance occurs at the end of S_3 . To build the balance evaluation model, we propose to extract the following features from the subject's CoP-CoM trajectory during the GI task.

- The maximum 2D CoP-CoM distance.
- The range of motion of the subject's CoM, in the x - and y -direction separately.

In our data collection process, each subject was required to perform three repetitions of GI on each leg. The motion of each subject (including all the six repetitions) constitutes a data sample. Therefore, there are $3 \times 6 = 18$ features in the input for each sample. Similar to the CoM estimation model, we propose a data augmentation approach for the balance evaluation model to create more training samples and avoid over-fitting. For the three repetitions that a subject performs on the left leg (e.g., L1, L2, L3), the order of the repetitions does not affect the overall performance of

the subject and the PT's evaluation. Therefore, the output of this sample should remain unchanged if the order of the three repetitions on the left leg is changed (e.g., L3, L1, L2). Based on the above insight, we propose the following data augmentation approach. For the three repetitions on the left leg, there are $3! = 6$ types of permutations. Similarly, there are six types of permutations for the three repetitions on the right leg. Therefore, we can generate $6 \times 6 = 36$ samples from each original sample by changing the order of the repetitions. We propose to train a Random Forest (RF) classifier [28] to estimate a balance level from the input features. During the data collection, the subject's balance ability was tested clinically by the PT with the mini-BESTest and used as the ground truth. The mini-BESTest scores were classified into four levels as follows.

- Level 4 (score 28): no balance problem, no fall risk
- Level 3 (score 18 ~ 27): mild balance problems, no fall risk.
- Level 2 (score 11 ~ 17): medium balance problems, medium fall risk.
- Level 1 (score 0 ~ 10): severe balance problems, high fall risk.

The balance level calculated from the PT score was used as the ground truth to train the balance evaluation model. The RF classifier takes all the 18 features as input and provides an estimate of the balance level as the output. Based on the study of Leddy *et al.* [31], patients with PD who get a score lower than 63% of the total score (i.e., $28 \times 63\% = 17.6$) on the mini-BESTest have fall risk. Therefore, Level 1 and 2 in our proposed balance evaluation system indicates fall risk. By using the proposed balance evaluation system, the patient is able to monitor his/her balance level and fall risk using a portable depth camera and WBB at home or any other place, which enables on-demand balance evaluation.

IV. RESULTS

In this section, we will first present our data collection process, then introduce the implementation details, finally evaluate the performance of the proposed CoM estimation and balance evaluation system.

A. DATA COLLECTION

This study was approved by the Institutional Review Board at University of California, San Diego (protocol #181413X). 41 subjects (age 23 ~ 81, 26 males, 15 females) participated in this study, including 21 healthy subjects and 20 patients with PD. To validate that our proposed model is able to learn the body parameters of the subject, we have recruited subjects of different body types (height 155 ~ 190 cm, weight 44 ~ 96 kg). All subjects signed the informed consent form. There were two stages in our data collection process. In the first stage, we collected data to train and test the proposed CoM estimation model. Each subject stood on the WBBs (shown in Figure 4) and performed the following static postures on four body parts.

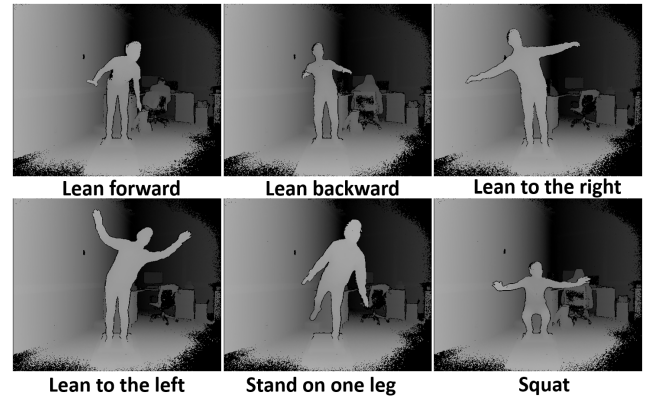


FIGURE 10. Examples of static postures collected in our experiments.

Trunk: keep it upright, or lean to left/right/front/back with different angles.

Legs: squat with different angles, stand on one leg.

Arms: different positions of the left and right arm.

Feet: different positions of the left and right foot.

Figure 10 shows some examples of the postures we have collected in our data collection.

The two WBBs recorded the CoP position, which was equivalent to the horizontal CoM position. We also used a Kinect sensor to capture the depth images of the subject. The WBB and the Kinect sensor were synchronized and the framerate was 30 frames per second. In the second stage, we collected data during the GI task for the balance evaluation system. Each subject stood on the WBB #1, made a step forward on the WBB #2 according to his/her natural walking, and steadily stepped off the board. Each subject performed three repetitions on the left and right leg separately. The CoP positions and depth images were also recorded by the WBB and the Kinect camera. The subject's dynamic balance was tested using the mini-BESTest by the PT as the ground truth.

B. IMPLEMENTATION DETAILS

For the CoM estimation model, we used $[-40, 40]$ (pixels), $[-0.2, 0.2]$ (depth value), and $[-15, 15]$ (pixels), for the random shift in the x -, y - (depth), and z -direction in the data augmentation. In the heatmap of the ground-truth CoM, we used Gaussian distribution with standard deviation of 3 and 2, in the x - and y -direction. There are five Conv units in the CNN-based model. In each Conv unit, 8, 16, 32, 64, 128 channels were used for the Conv layer respectively. The number of channels was selected empirically. The BN momentum was set to 0.9. When training the model, we used an Adam optimizer [29] to minimize the cross-entropy loss. The batch size was 64 and the learning rate was $5e-4$. For the proposed coarse-to-fine approach, we trained three models using $DI_1 = 8mm$, $DI_2 = 4mm$, and $DI_3 = 2mm$. For the balance evaluation model, we trained a RF classifier with 300 trees in the forest. The input of the classifier is 18-dimensional and the output is four categories. We used Gini impurity to measure the quality of a split when constructing the trees.

TABLE 1. CoM estimation error and requirements of each method.

Method		Error (mm)		Requirements
		x	y	
SESC [20]		17	23	Motion capture sensor. Identification needed for each new subject.
Voxel reconstruction [12]		8 ~ 15		Five cameras. Camera calibration and synchronization.
CNN-based (our previous work) [19]	Existing subjects	6.0	9.3	Single depth camera. No calibration or identification needed.
	New subject	8.9	17.2	
CNN + coarse-to-fine (proposed)	Existing subjects	5.2	8.4	
	New subject	7.8	15.7	

C. CoM ESTIMATION RESULTS

To validate the proposed CoM estimation model, we calculate the estimation error as the distance between the ground-truth CoM position and the estimated position (i.e., the center of the output class). Firstly we validate the performance of the model on existing subjects. We randomly split all the samples into three parts: a training set (including 64% of the samples), a validation set (including 16% of the samples), and a test set (including the rest 20% of the samples). Secondly we validate the performance of the model on a new subject. The samples of 40 subjects are used for training and validation and the samples from the 41st subject are used for testing. This process is repeated for 10 times and the average results are presented. We compare the results of the following methods: the CNN-based model proposed in our preliminary work [19], the CNN + coarse-to-fine approach proposed in this paper, and two state-of-the-art methods: the SESC method [20] and the voxel reconstruction method [12]. Table 1 presents the estimation error and the requirements of each method.

When testing on existing subjects, the proposed CNN-based method (proposed in [19] and in this paper) achieves the lowest estimation error. When testing on a new subject, the estimation error achieved by our methods increases a little bit, but still outperforms the SESC method in both x - and y - directions. In additional, the identification phase required by the SESC method is not convenient for home and clinical use. For example, an existing subject may need to go through the identification phase again if he/she gains or loses weight. In comparison, the proposed CNN-based approach is able to learn the subject's body parameters from the depth image without any identification process. Compared with the voxel reconstruction method [12], our proposed approach achieves comparable accuracy results, but requires only a single depth camera and avoids the complicated calibration and synchronization among multiple cameras. Therefore, it is more convenient for home and clinical use. Comparing the estimation error in the x - and y -direction, we can see that the error in the y -direction (depth direction) is higher, which

TABLE 2. Comparison of the training and inference times.

Feature		Total training time (s)	Average inference time (ms)
Joint heatmaps [19]		2252.1	38.0
Colored skeleton image (proposed)	Single model	322.7	12.6
	Coarse-to-fine	963.5	38.6

is due to the fact that the back side of the body cannot be captured by the single depth camera.

Moreover, the coarse-to-fine approach proposed in this paper further reduces the estimation error by about 10%, compared with the preliminary model in [19]. Besides, Table 2 shows the comparison of the total training time (i.e., the total time to update the parameters in one epoch) and the average inference time (i.e., the average time on each sample) by using the proposed colored skeleton image (discussed in Section III-B) and the joint heatmaps proposed in [19]. For the colored skeleton image approach, we show the training and inference times using single model and multiple models in the coarse-to-fine approach. The running time is tested on an Intel Xeon E5-1650 CPU and an NVIDIA GeForce GTX 1080 Ti GPU. We can see that the training and inference times of single model are significantly reduced by using the proposed colored skeleton image in the input of the CNN model. Although the proposed coarse-to-fine approach increases the training and inference times by using multiple models, it can still achieve much less training time and comparable inference time compared with the preliminary model proposed in [19], while significantly reducing the estimation error (see Table 1). Therefore, it can be concluded that the CoM estimation model proposed in this paper improves our preliminary model proposed in [19] by significantly reducing the estimation error, as well as the training and inference times.

D. BALANCE EVALUATION RESULTS

To show the performance of the proposed balance evaluation system, we first provide more details on the collected data during the GI task. Table 3 shows the average value of each input feature (discussed in Section III-C) for each balance level. We can see that subjects in lower balance level (i.e., worse balance) have smaller CoP-CoM distance. Similarly, subjects with worse balance also show smaller range of motion in their CoM position in the y -direction (i.e. the anterior-posterior direction), which indicates that subjects with worse balance have smaller step length and smaller body movement during the GI task. For the range of motion in the x -direction (i.e., the medio-lateral direction), subjects in level 4 (who got full score 28 in the mini-BESTest) have higher range of motion. However, there is no significant trend for the other three levels.

To validate the proposed RF-based balance evaluation model, we conduct experiments using 10-fold cross validation, with 90% of the collected samples used for training and

TABLE 3. Average feature values for each balance level.

Feature		Balance level			
		1	2	3	4
Maximum CoP-CoM distance (mm)		137.3	177.3	217.8	264.2
Range of motion of the CoM position (mm)	x	22.1	18.4	15.1	50.0
	y	86.3	115.8	178.7	74.1

TABLE 4. Sensitivity and specificity using the proposed balance evaluation model.

	Balance level				Fall risk	
	1	2	3	4	Yes	No
Sensitivity	87.5%	82.5%	82.0%	85.0%	93.8%	95.4%
Specificity	98.8%	93.5%	89.4%	97.7%	95.4%	93.8%

10% for testing. The proposed data augmentation approach is applied to the training samples. We calculate the sensitivity (i.e., the proportion of actual positive samples that are correctly classified) and specificity (i.e., the proportion of actual negative samples that are correctly classified) for each level and report the results in Table 4. We also show the results on the two categories: with fall risk (levels 1 and 2) and without fall risk (i.e., levels 3 and 4). We can see that the proposed RF-based model can achieve high sensitivity and specificity for the four levels ($>80\%$) and the two categories ($>90\%$). Besides, all the classification error is only one level (i.e., no sample is misclassified as a level higher or lower than the ground-truth level by two levels or more). Therefore, it can be concluded that the proposed balance evaluation system is able to provide accurate and quantitative balance assessments like a human PT. The high accuracy also demonstrates that the proposed CoM estimation model works for dynamic postures. By using the proposed balance evaluation system, the patient can measure his/her balance level using a simple GI task at home or in the clinic. The quantitative balance level can help the patient (and his/her PT) evaluate progress in physical therapy training, select the proper training programs, and predict the fall risk.

V. CONCLUSION AND FUTURE WORK

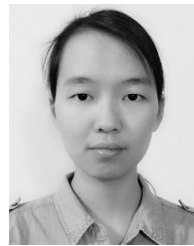
In this paper, we propose a balance evaluation system using camera and WBB sensors to enable on-demand balance evaluation for home and clinic-based physical therapy. To develop this system, we first propose a CoM estimation model to estimate the CoM position of the human body from a depth image. Experimental results on the CoM estimation model demonstrate its superiority over other CoM estimation techniques, including high accuracy and the ease-of-use. Based on the CoM estimation model, we further propose the balance evaluation system to estimate a quantitative balance level from the subject's performance during a GI task. Experimental results show that the proposed model can accurately estimate a balance level that is consistent with the human PT's evaluation in traditional balance tests. By using portable and inexpensive sensors, the proposed balance evaluation system enables on-demand balance evaluation for home and clinical use and has the potential of significantly reducing clinic visit requirements and reducing cost for both the patients and care providers.

For future work, we would like to improve the accuracy of the proposed CoM estimation model, especially in the depth direction. In our current experiments, the WBBs were placed in front of the camera so only the front view was captured. In the future, we would like to capture different views of the user body. Moreover, we would like to test the accuracy of the WBB by comparing it with a laboratory-grade force plate in our data collection. We also plan to improve the current balance evaluation system to provide more detailed balance assessments (e.g., continuous balance scores) instead of the four levels. Besides, the GI task discussed in this paper may be limited for balance evaluation. We plan to explore more training exercises in physical therapy to achieve more comprehensive balance evaluation for patients with balance problems.

REFERENCES

- [1] M. Mak and M. Auyeung, "The mini-BESTest can predict parkinsonian recurrent fallers: A 6-month prospective study," *J. Rehabil. Med.*, vol. 45, no. 6, pp. 565–571, 2013.
- [2] K. O. Berg, S. L. Wood-Dauphinee, J. I. Williams, and B. Maki, "Measuring balance in the elderly: Validation of an instrument," *Can. J. Public Health = Revue Canadienne De Sante Publique*, vol. 83, no. 2, pp. 7–11, 1992.
- [3] F. Franchignoni, F. Horak, M. Godi, A. Nardone, and A. Giordano, "Using psychometric techniques to improve the Balance evaluation systems test: The mini-BESTest," *J. Rehabil. Med.*, vol. 42, no. 4, 2010, Art. no. 323331.
- [4] A. K. Mishra, M. Skubic, B. W. Willis, T. Guess, S. S. Razu, C. Abbott, and A. D. Gray, "Examining methods to estimate static body sway from the kinect V2.0 skeletal data: Implications for clinical rehabilitation," in *Proc. 11th EAI Int. Conf. Pervas. Comput. Technol. Healthcare PervasiveHealth*, 2017, pp. 127–135.
- [5] M. W. Kennedy, J. P. Schmiedeler, C. R. Crowell, M. Villano, A. D. Striegel, and J. Kuitse, "Enhanced feedback in balance rehabilitation using the nintendo wii balance board," in *Proc. IEEE 13th Int. Conf. e-Health Netw., Appl. Services*, Jun. 2011, pp. 162–168.
- [6] H. L. Bartlett, L. H. Ting, and J. T. Bingham, "Accuracy of force and center of pressure measures of the wii balance board," *Gait Posture*, vol. 39, no. 1, pp. 224–228, Jan. 2014.
- [7] W. Young, S. Ferguson, S. Brault, and C. Craig, "Assessing and training standing balance in older adults: A novel approach using the 'Nintendo Wii' balance board," *Gait Posture*, vol. 33, no. 2, pp. 303–305, Feb. 2011.
- [8] J. D. Holmes, M. E. Jenkins, A. M. Johnson, M. A. Hunt, and R. A. Clark, "Validity of the Nintendo Wii balance board for the assessment of standing balance in Parkinson's disease," *Clin. Rehabil.*, vol. 27, no. 4, pp. 361–366, 2013.
- [9] D. A. Winter, *Biomechanics and Motor Control of Human Movement*. Hoboken, NJ, USA: Wiley, 2009.
- [10] S.-C. Chen, H.-J. Hsieh, T.-W. Lu, and C.-H. Tseng, "A method for estimating subject-specific body segment inertial parameters in human movement analysis," *Gait Posture*, vol. 33, no. 4, pp. 695–700, Apr. 2011.
- [11] S. Cotton, A. P. Murray, and P. Fraisse, "Estimation of the center of mass: From humanoid robots to human beings," *IEEE/ASME Trans. Mechatronics*, vol. 14, no. 6, pp. 707–712, Dec. 2009.
- [12] T. Kaichi, S. Mori, H. Saito, K. Takahashi, D. Mikami, M. Isogawa, and H. Kimata, "Estimation of center of mass for sports scene using weighted visual hull," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 1809–1815.
- [13] S. Ebrahimi Kahou, V. Michalski, K. Konda, R. Memisevic, and C. Pal, "Recurrent neural networks for emotion recognition in video," in *Proc. ACM Int. Conf. Multimodal Interact. ICMI*, 2015, pp. 467–474.
- [14] R. Alazrai, Y. Mowafi, and E. Hamad, "A fall prediction methodology for elderly based on a depth camera," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 4990–4993.
- [15] J. Shotton, R. Girshick, A. Fitzgibbon, T. Sharp, M. Cook, M. Finocchio, R. Moore, P. Kohli, A. Criminisi, A. Kipman, and A. Blake, "Efficient human pose estimation from single depth images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 12, pp. 2821–2840, Dec. 2013.

- [16] W. Wei, Y. Lu, E. Rhoden, and S. Dey, "User performance evaluation and real-time guidance in cloud-based physical therapy monitoring and guidance system," *Multimedia Tools Appl.*, vol. 78, no. 7, pp. 9051–9081, Apr. 2019.
- [17] W. Wei, C. McElroy, and S. Dey, "Towards on-demand virtual physical therapist: Machine learning-based patient action understanding, assessment and task recommendation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 9, pp. 1824–1835, Sep. 2019.
- [18] C. J. Hass, D. E. Waddell, R. P. Fleming, J. L. Juncos, and R. J. Gregor, "Gait initiation and dynamic balance control in parkinson's disease," *Arch. Phys. Med. Rehabil.*, vol. 86, no. 11, pp. 2172–2176, 2005.
- [19] W. Wei and S. Dey, "Center of mass estimation for balance evaluation using convolutional neural networks," in *Proc. IEEE Int. Conf. Healthcare Informat. (ICHI)*, Jun. 2019, pp. 1–7.
- [20] A. González, M. Hayashibe, V. Bonnet, and P. Fraisse, "Whole body center of mass estimation with portable sensors: Using the statically equivalent serial chain and a kinect," *Sensors*, vol. 14, no. 9, pp. 16955–16971, Sep. 2014.
- [21] J. Swanenburg, E. D. de Bruin, K. Favero, D. Uebelhart, and T. Mulder, "The reliability of postural balance measures in single and dual tasking in elderly fallers and non-fallers," *BMC Musculoskeletal Disorders*, vol. 9, no. 1, p. 162, Dec. 2008.
- [22] *Kinect*. Accessed: Aug. 5, 2019. [Online]. Available: <https://www.xbox.com/en-US/kinect>
- [23] A. Huurnink, D. P. Franz, I. Kingma, and J. H. van Dieën, "Comparison of a laboratory grade force platform with a nintendo wii balance board on measurement of postural control in single-leg stance balance tasks," *J. Biomech.*, vol. 46, no. 7, pp. 1392–1395, Apr. 2013.
- [24] J. J. Tompson, A. Jain, Y. LeCun, and C. Bregler, "Joint training of a convolutional network and a graphical model for human pose estimation," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, Montreal, QC, Canada, Dec. 2014, pp. 1799–1807.
- [25] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *Handbook Brain Theory Neural Netw.*, vol. 3361, no. 10, p. 1995, 1995.
- [26] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, Lake Tahoe, America, Dec. 2012, pp. 1097–1105.
- [27] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: <http://arxiv.org/abs/1502.03167>
- [28] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [29] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: <http://arxiv.org/abs/1412.6980>
- [30] S. E. Halliday, D. A. Winter, J. S. Frank, A. E. Patla, and F. Prince, "The initiation of gait in young, elderly, and Parkinson's disease subjects," *Gait Posture*, vol. 8, no. 1, pp. 8–14, Aug. 1998.
- [31] A. L. Leddy, B. E. Crouner, and G. M. Earhart, "Utility of the mini-BESTest, BESTest, and BESTest sections for balance assessments in individuals with parkinson disease," *J. Neurologic Phys. Therapy*, vol. 35, no. 2, p. 90, 2011.
- [32] Y.-L. Hsu, P.-C. Chung, W.-H. Wang, M.-C. Pai, C.-Y. Wang, C.-W. Lin, H.-L. Wu, and J.-S. Wang, "Gait and balance analysis for patients with Alzheimer's disease using an Inertial-Sensor-Based wearable instrument," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 6, pp. 1822–1830, Nov. 2014.
- [33] P. Esser, H. Dawes, J. Collett, and K. Howells, "IMU: Inertial sensing of vertical CoM movement," *J. Biomech.*, vol. 42, no. 10, pp. 1578–1581, Jul. 2009.
- [34] B. Fasel, J. Spörri, P. Schütz, S. Lorenzetti, and K. Aminian, "An inertial sensor-based method for estimating the Athlete's relative joint center positions and center of mass kinematics in alpine ski racing," *Frontiers Physiol.*, vol. 8, p. 850, Nov. 2017.
- [35] G. Pavlakos, X. Zhou, K. G. Derpanis, and K. Daniilidis, "Coarse-to-fine volumetric prediction for single-image 3D human pose," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 7025–7034.
- [36] A. Newell, K. Yang, and J. Deng, "Stacked hourglass networks for human pose estimation," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Amsterdam, The Netherlands, Oct. 2016, pp. 483–499.



WENCHUAN WEI (Member, IEEE) received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2014. She is currently pursuing the Ph.D. degree with the University of California San Diego, La Jolla, CA, USA. Her research interests include digital health, machine learning, and multimedia.



CARTER MCELROY received the master's degree in physical therapy from Northern Arizona University. He is currently a Physical Therapist who specializes in treating people with various balance conditions. He often works with patients who have Parkinson's disease or vestibular disorders, as well as stroke survivors. He is also experienced in treating patients with various orthopedic injuries. He teaches individuals how to prevent or manage their condition so they gain improved understanding of their movement deficits to improve their function. Previously, he was an original member of the Movement Disorder Team, UC San Diego Health, where he has spoken at Community Symposiums on the topic of balance and mobility related to Parkinson's disease. He has also participated in research and published in the area of Parkinson's disease.



SUJIT DEY (Fellow, IEEE) received the Ph.D. degree in computer science from Duke University, in 1991.

He is currently a Professor with the Department of Electrical and Computer Engineering, the Director of the Center for Wireless Communications, and the Director of the Institute for the Global Entrepreneur, University of California, San Diego. He heads the Mobile Systems Design Laboratory, developing innovative and sustainable edge computing, networking and communications, multi-modal sensor fusion, and deep learning algorithms and architectures to enable predictive personalized health, immersive multimedia, and smart transportation applications. He has created inter-disciplinary programs involving multiple UCSD schools as well as community, city and industry partners, notably the Connected Health Program, in 2016, and the Smart Transportation Innovation Program, in 2018. In 2017, he was appointed as an Adjunct Professor with the Rady School of Management, and the Jacobs Family Endowed Chair in Engineering Management Leadership. He has served as the Faculty Director of the von Liebig Entrepreneurism Center, from 2013 to 2015, and the Chief Scientist of mobile networks at Allot Communications, from 2012 to 2013. In 2015, he co-founded igenEnergi, providing intelligent battery technology and solutions for EV mobility services. He founded Ortiva Wireless, in 2004, where he served as its founding CEO and later as a CTO and a Chief Technologist till its acquisition by Allot Communications, in 2012. Prior to Ortiva, he served as the Chair of the Advisory Board of Zyrray Wireless till its acquisition by Broadcom, in 2004, and an advisor to multiple companies including ST Microelectronics and NEC. Prior to joining UCSD in 1997, he was a Senior Research Staff Member at NEC C&C Research Laboratories in Princeton, NJ. He has coauthored more than 250 publications, and a book on low-power design. He holds 18 U.S. and two international patents, resulting in multiple technology licensing and commercialization. He was a recipient of nine IEEE/ACM Best Paper Awards, and has chaired multiple IEEE conferences and workshops.

...