IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 2, 2019, accepted July 24, 2019, date of publication July 30, 2019, date of current version August 15, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2932101

Physiological Function Assessment Based on Kinect V2

WENMING CAO[1,2], JIANQI ZHONG[1], GUITAO CAO[3], AND ZHIQUAN HE[1]

[1]Guangdong Multimedia Information Service Engineering Technology Research Center, Shenzhen University, Shenzhen 518060, China
[2]Video Processing and Communication Lab, Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO 65211, USA
[3]School of Computer Science and Software Engineering, East China Normal University, Shanghai 200062, China

Corresponding author: Zhiquan He (zhiquan@szu.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 61771322, Grant 61871186, and Grant 61375015, in part by the China Scholarship Council (CSC), and in part by the Fundamental Research Foundation of Shenzhen under Grant JCYJ20160307154630057.

ABSTRACT This paper presented a framework based on fog computing for convenient and efficient Physiological Function Assessment, which consists of three parts: 1)measuring the degree of joint mobility; 2) investigating the abnormality of actions of upper limbs; and 3) abnormal gait detection for lower limbs. Especially, we introduced semi-automatic Rapid Upper Limb Assessment (RULA) using Kinect v2 for the upper limb motion evaluation. Since a specific action can be described by action sequences of different length, we used dynamic time warping (DTW) to find the similarity between action sequences with different length. Traditional DTW algorithm does not work well when the action sequences are long and complex. To address this problem, we improved the DTW method by modifying the mapping relationship and limiting the computation space. Our modified DTW algorithm was evaluated on a standard 3D action dataset (SYSU 3D HOI) and Human Upper Action dataset (HUA), achieving the accuracy of 83.75%, 89.50%, respectively. The result is significantly better than the traditional DTW and the reported methods. In our previous work, we described the framework and how to make physiological function assessment. The goal of this paper is to 1) enrich the experiments of previous work and 2) introduce the framework of using RULA for physiological function assessment. All the tests have been done in this framework based on fog computing.

INDEX TERMS Physiological function assessment, fog computing, kinect, human activity recognition, RULA.

I. INTRODUCTION

A. PHYSIOLOGICAL FUNCTION ASSESSMENT FRAMEWORK BASED ON FOG COMPUTING

At present, the commonly used methods of physiological function testing mainly include photographic method, wearable device method and other artificial detection methods. These tests need to be completed with the help of professional staffs in specific occasions, which are expensive and time-consuming. In order to conduct Physiological Function Assessment(PFA) in hospital even at home efficiently and conveniently, many researchers have proposed to use cloud computing to solve this problem. A remote monitoring platform based on cloud computing was proposed to monitor the health status for the elderly in real time [1]. According to monitor data, doctors can give suggestions for better treatment in details. The authors presented a real-time monitoring method of multi-physiological parameters cluster applying to the remote monitoring platform for efficiently detection and evaluation [2]. As mobile Internet and Internet of Things technologies develop further, cloud computing is becoming more and more accessible. However, a study by Gartner suggested that centralizing all this distributed data into one data center in the big data era is technically and economically infeasible [3]. Firstly, the growth rate of data exceeds the growth rate of network bandwidth, which will cause the cloud server to be overloaded, causing network congestion, and users' requests are unable to get response in time. Secondly, the physical distance between the cloud server and the terminal device is generally far, which leads to the increasing delay of data transmission between them.

Hence, this study presented a new framework based on fog computing to reduce the problem caused by cloud computing. Between the cloud server and the terminal device,

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

some special equipment is used as fog service for computing, storage. Data processing is scattered on fog servers located on the edge of the network, so that the data processing can be close to the terminal equipment and offer service to the users. In this framework, data would be sent to the fog servers for the priority processing, which could reduce the unnecessary network transmission and increases the response speed of application systems [4].

B. PFA BASED ON SKELETON TRACKING

Human movement status is key to physiological function assessment, which is also a main indicator of human health. In this paper, we focused on evaluation for human body movement for PFA. The decline in the function of human movement, which is affected by joints and muscles of human body, will lead to many diseases, such as Periarthritis [5], Osteoarthritis [6], Osteoporosis [7], Stroke [8] and even paralysis. Therefore, PFA for the elderly in time, and even monitoring and evaluating them over time, will enable early diagnosis of diseases and help find the best treatment.

Many researcher have explored different aspects for human body function, which mainly include 1) measuring the Degree of Joint Mobility (DJM), 2) investigating the abnormality of actions of upper limbs and 3) abnormal gait detection for lower limbs. Hu et al. calculated DJM by using a NDI Polaris Spectra to track the 3D position of the reflective ball on the limb bones [9]. Kinect device was used to monitor and detect elderly people when they are likely to fall by measuring their gaits in [10]. Hu et al. used Kinect mounted on the pedals to get walkers' leg information (angles of lower joints, relative distance between joints) to carry out gait detection [11]. Standen et al. used Kinect to capture the scene and then tested the posture of the upper limbs [12].

These methods obtain the locations of body joints and evaluate the body movements. The Kinect v2, released in 2013, use a technology of time-of-flight, has a excellent performance of tracking the positions of 25 body joints. Kinect is widely used in the study of physiological function assessment. Li et al. used Kinect to improve the efficiency and accuracy of rapid upper limb assessment (RULA) while reducing the workload [13]. Chen et al. designed a physiological function improvement game for old people based on Kinect system [14]. Bernardino et al. used Kinect to collect a dataset, which contains a number of fitness tests according to the physiological function standards of reliability and validity [15].

Joints status is an important indicator of physiological functions, which is reflected by the action of the human body such as hand waving, TW stretching etc. Therefore, the evaluation for upper limbs can be done well with the pose being recognized. A typical way for action recognition is to search the action sequence database finding the similar sequence with smallest distance. A specific action can be described by an action sequence which consists of extracted features. People perform actions differently, even for the same action type, e.g. walking. Therefore, the action sequences describing the same action could have different length. To compare two action sequences, Dynamic Time Warping (DTW) can be used to find the smallest distance. DTW has been reported in several publications [16]–[18], where it has been shown by the preliminary test on Japanese digit words indicating DTW's effectiveness. However, the matching process is time-consuming and inaccurate in some cases, especially when the two sequences are complicated. In order to improve the matching performance, we modified the DTW algorithm [19] with a better distance measure and a limited the search space during the dynamic programming (dp) procedure.

Rapid Upper Limb Assessment(RULA) [20], is one of the most popular observational methods [21] used to evaluate the risk of work-related musculoskeletal disorders (WMSDs). This weakness of RULA is the requirement of a field expert to analyze the postures, which is time-consuming and laborsome. Many methods have been developed for this problem, suggesting a semi-automatic approaches using low-cost and easy-to-use cameras. Several work explored the accuracy of kinematic data provided by Kinect v1 in various applications [22]–[25]. Kinect v1 was used to make the RULA for posture analysis, which shows the advantages: calculating in real time, minimizing the time consumption of the assessing procedures, reducing the bias from the analyst [26]. Kinect v2 was also used on the validation of RULA grandscores. Manghisi et al. compared the RULA scores from Kinect v2 with the ones obtained by the RULA experts, Jack-TAT and optical motion capture system [27]. The results above show that Kinect v2 device is an effective tool for RULA analysis. Therefore, we introduced this method into our PFA framework and made posture evaluation test based on it. In addition, instead of setting the neck twist manually as in [27], we used the joint orientation data provided by Kinect to calculate the neck twist in real time.

Gait abnormal detection is the key to gait analysis. We have modified the DTW and used K-Nearest Neighbor (KNN) [28] to process action sequences detecting whether the movement is abnormal as in [29].

In our previous work [30], we have introduced the fog computing framework for PFA, abnormality assessment and modified DTW algorithm. Based on this, in this paper, we conducted more experiments with more subjects to validate the effectiveness of joint angle measurements, motion recognition and gait abnormality detection with the modified DTW (MDTW) algorithm, and introduced the semi-automatic approach for RULA assessment using Kinect sensor.

The paper is organized as follows. Section II describes the framework of PFA based on fog computing, and how to make PFA based on Kinect v2. In Section III, we validated the effectiveness of the proposed method for motion recognition. And, we compared the grand-scores from Kinect v2 with a RULA expert and validate the feasibility of the RULA using Kinect in our PFA framework. Section IV provides conclusions for this paper and suggestions for future studies.

II. METHOD

A. PFA BASED ON FOG COMPUTING

The system based on fog computing consists of four layers: 1) fog servers, 2) terminal users, 3) cloud servers and 4) medical center (Fig. 1). The layer of terminal users collect physiological data, which would be transferred to fog server immediately. Fog servers analyze these data and return a report about these physiological data, e.g. the status of the elderly. The analysis process is based on the cached database in the hospital or medical center. In order to adapt to different scenarios, the fog server would cache lots of useful information for users according to the characteristics of the application scenes [31]. For instance, in hospitals, fog server should save standard biometric data to help doctor judge whether there is something wrong with the elderly. The report given by fog servers would be fed back to terminal users and medical center. Based on the feedback, the doctor can give timely advise to the terminal users. Given that the primary application of the system is to carry out physiological function assessment efficiently, the cloud server would take over the works above when the data to process is too complex to handle for fog servers.

Description of Terminal Users: This layer can have many electronic sensors, like cameras and wearable sensors, which can capture the physiological state of the elderly. In our study, Kinect v2 sensor was used as data collecting tool.

Description of Fog Computing Layer: The fog computing layer consists of various small computing devices, routers, small servers, switches, etc., which are at the edge of the network and close to the user. Therefore, the data from terminal users could be sent to fog server instead of remote cloud server for the priority processing. As illustrated in Fig. 1, fog computing layer analyzes the physiological data from terminal users, searches the database, and then feeds back to terminal users as well as medical centers. To ensure the repeatability and reproducibility, the fog servers may update its database using the new results to make it more complete and accurate.

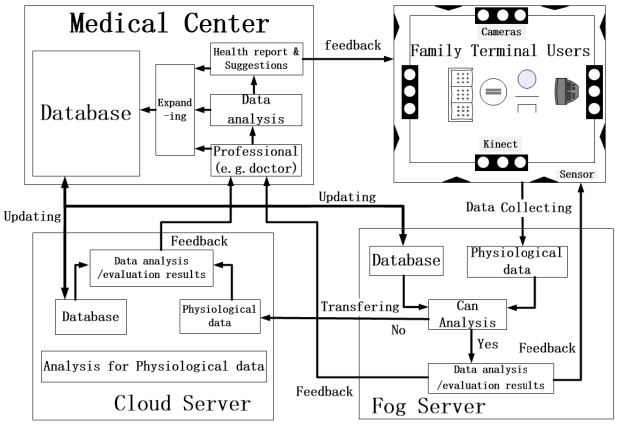

FIGURE 1. Family physiological function assessment based on fog computing.

Description of Cloud Computing Layer: Cloud servers are usually powerful in computing and have large storage. When a task needs to handle huge amount of data, the data will be transferred from the fog servers to the cloud servers. As illustrated in Fig. 1, the cloud servers process the data that the fog server cannot analyze and return a feedback results to the medical centers as well as terminal users.

Description of Medical Center: Feedback resulted from fog or cloud servers would serve as a reference for doctors' diagnosis. Doctors would give suggestions to the family according to these feedback. In addition, the valuable data in these feedback are used to enrich the database, which can be a reference to the analysis of fog or cloud servers.

B. PFA WITH KINECT V2

In this work, we mainly focused on the movement assessment using RGB-D camera (Microsoft Kinect v2) for physiological function assessment. Kinect v2 can track the locations of 25 major nodes of the human body (Fig. 2) and present them in the body image, which lays the foundation of motion detection [32], [33]. The Kinect coordinate system is defined as the right-handed coordinate system, in which the Z-axis is pointing in the direction as the camera sees and Y-axis pointing upward. We studied three aspects: degree of joint mobility (DJM), action recognition, and abnormal gait detection to assess whether the movements are normal.

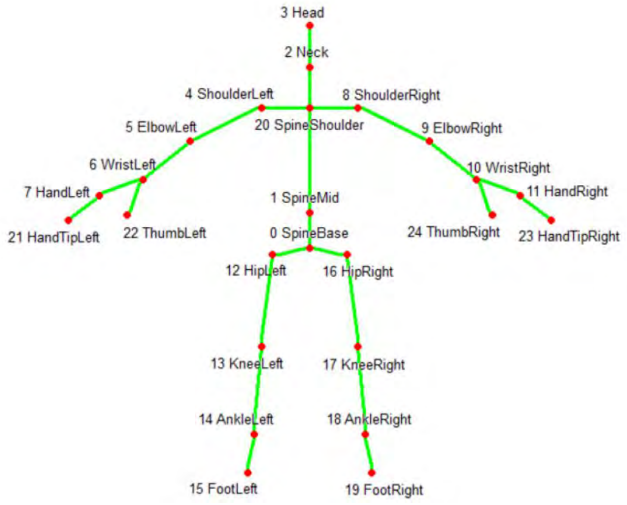

FIGURE 2. The 25 body joints Kinect v2 can capture.

1) DEGREE OF JOINT MOBILITY (DJM)

DJM is defined as the angle range of body joints, e.g. the extension and flexion of joints, which reflects a person's health condition. Some diseases such as muscle weakness, frozen shoulder and osteoarthritis caused by joint or muscle damage would lead to the abnormality of DJM. Therefore, the investigation of DJM is essential for PFA. According to the range of the joint mobility, we can tell if the joint movement is normal. In order to assess physiological function

TABLE 1. The chart for DJM in unit of degree of five main joints of human body. The scores in the first row correspond to the maximum angle range the joint can reach.

Score	1	2	3	4
Knee extension/flexion	$45 \sim 90$	$90 \sim 150$	$0 \sim 45$	$150 \sim 180$
Hip extension/flexion	$0 \sim 15$	$15 \sim 30$	$30 \sim 45$	$45 \sim 90$
shoulder extension/flexion	$0 \sim 60$	$60 \sim 120$	$120 \sim 150$	$150 \sim 180$
Elbow extension/flexion	$45 \sim 90$	$90 \sim 150$	$0 \sim 5$	$150 \sim 180$
Neck extension/flexion	$-15 \sim 15$	$-45 \sim 45$	$-60 \sim 60$	$-90 \sim 90$

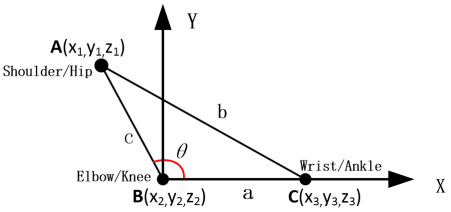

FIGURE 3. Joint angles for elbow or knee. A is the shoulder/hip joint, B is the elbow/knee joint, and C is the wrist/ankle joint. The coordinates of A, B and C are given by Kinect.

in detail, we created a chart to describe the joint mobility (see Table 1). We studied five main joints involved in daily activities of human body: neck, shoulder, elbow, hip and knee. The scores of the first row in the table are determined by the joint motion range. The smaller action range of the joint, the lower the score. For example, for the hip joint, if the action range is less than 15°, the score is 1, but if the joint can do an action of wider range of 45° to 90°, the score is 4.

a: ELBOW/KNEE EXTENSION AND FLEXION

For the elbow/knee joints, the corresponding joint angle can be obtained easily by calculating the angle between $\overrightarrow{BC}$ and $\overrightarrow{BA}$ (Fig. 3). Let $A(x_1, y_1, z_1)$ denote the shoulder/hip joint, $B(x_2, y_2, z_2)$ the elbow/knee joint, and $C(x_3, y_3, z_3)$ the wrist/ankle joint. When the tester moves, the angle between $\overrightarrow{BA}$ and $\overrightarrow{BC}$ ($\angle ABC$) changes. The distance between point $M(x_1, y_1, z_1)$ and $N(x_2, y_2, z_2)$ is defined as

$$d(M, N) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (1)$$

where M and N can be any of A, B and C. The joint angle is

$$\theta = arccos\frac{a^2 + c^2 - b^2}{2ac} \quad (2)$$

where $a = d(B, C)$, $b = d(A, C)$ and $c = d(A, B)$.

b: NECK EXTENSION AND FLEXION

We assessed the neck extension/flexion by computing the angle θ between the vector connecting the neck to the head and the vector connecting the SpineShoulder joint (Kinect v2 nomenclature from Fig. 2) to the neck. Fig. 4 shows a part of the skeleton information Kinect v2 captured. $A(x_1, y_1, z_1)$ denotes the SpineShoulder joint, $B(x_2, y_2, z_2)$ the Neck joint, and $C(x_3, y_3, z_3)$ the Head. The joint angle for the Neck can

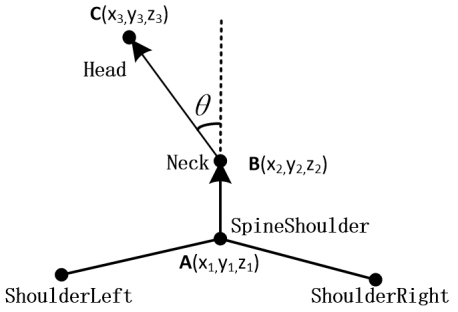

FIGURE 4. Neck extension or flexion calculation with the coordinates given by Kinect.

be calculated by

$$\theta = arccos\frac{\overrightarrow{AB} \bullet \overrightarrow{BC}}{\left|\overrightarrow{AB}\right| \left|\overrightarrow{BC}\right|} \quad (3)$$

where $\overrightarrow{AB} \bullet \overrightarrow{BC}$ is the dot product between $\overrightarrow{AB}$ and $\overrightarrow{BC}$. θ is positive when the Head is tilted to the ShoulderLeft, and negative when the Head is tilted to the ShoulderRight.

c: SHOULDER/HIP EXTENSION AND FLEXION

In Fig. 5, we defined the trunk vector ($\overrightarrow{CE}$) as the vector connecting the SpineShoulder (from Kinect v2 nomenclature) to the Spinebase. For shoulder and hip, no explicit angle can be used to describe the corresponding DJM effectively, so we introduced the angle between the trunk vector and the vector corresponding to the projection of $\overrightarrow{BA}$ or $\overrightarrow{FG}$ on the sagittal plane. The latter is evaluated as the one passing through the trunk vector and perpendicular to the straight line connecting the shoulder or knee. As showed in Fig. 5, θ can be obtained by calculating the angle between the trunk $\overrightarrow{CE}$ and $\overrightarrow{ba}$. It is worth noting that:

$$\overrightarrow{ba} = (\overrightarrow{EF} \times \overrightarrow{BA}) \times \overrightarrow{EF} \quad (4)$$

where $\overrightarrow{EF}$ (connecting the SpineBase joint and HipRight joint) represents the normal vector of the sagittal plane. $\overrightarrow{BA}$

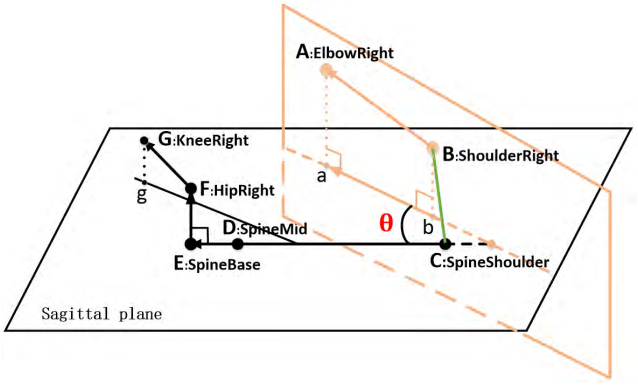

FIGURE 5. The joint of Shoulder and Hip assessment geometrical construction.

is the vector connecting shoulder to elbow. We used cross product to obtain $\overrightarrow{ba}$, θ can be computed by substituting Eq. (4) in (3).

2) ACTION RECOGNITION

Upper limbs motions are mainly expressed by the limbs and the trunk. Therefore, in this work, we extracted 7 joint points (shoulder, elbow,wrist and spineMid) to construct action feature (AF) in the upper part of human body. An AF consists of two components 1) the angle of joints and 2) the relative displacement between joints. Relative joint displacement is the spatial displacements of human skeletal joints based on the difference to the reference joints [34]. We calculated the angle of shoulder and elbow (4 angles including left and right), the relative position from wrist and elbow to the SpineMid respectively (4 distances including left and right). Therefore, we have a feature vector of 8 dimensions for each action.

$$AF = \{A_1, A_2, A_3, A_4, d_1, d_2, d_3, d_4\} \tag{5}$$

where A_i denotes the angle of joints, d_i is the distance between joints.

a: DTW ALGORITHM

In order to measure the similarity between two action sequences, DTW is used to find the alignment with smallest distance. Timing difference between two action sequences are eliminated by warping the time axis of one so that the maximum coincidence is attained with the other. Then, the time-normalized distance is calculated as the minimized residual distance between them. Suppose A and B are two action sequences

$$\begin{aligned} A &= \{a_1, a_2, ..., a_m\} \\ B &= \{b_1, b_2, ..., b_n\} \end{aligned} \tag{6}$$

If $m = n$, we can directly calculate the distance or similarity between A and B. If $m \neq n$, we set up an m×n score matrix (Fig. 6), where DTW algorithm assigns a score $D(i, j)$ to each matrix element by:

$$D(i,j) = min\{D(i,j-1), D(i-1,j), D(i-1,j-1)\} + d(i,j)$$

where $d(a_i, b_j)$ is the Euclidean distance $d(a_i, b_j) = (a_i - b_j)^2$. Then, try to find the optimal alignment path (the red path in Fig. 6) by

$$W = (w_1, ..., w_{k-1}, w_k), \quad k \in [max(m,n), m+n-1] \tag{7}$$

The optimal alignment minimizes

$$DTW(A,B) = \frac{\sqrt{\sum_{k=1}^{K} w_k}}{K} \tag{8}$$

where DTW(A,B) is considered as the distance between A and B. K is the length of the alignment path and w_k is the mapping between a_i and b_j at the k-th node of the path.

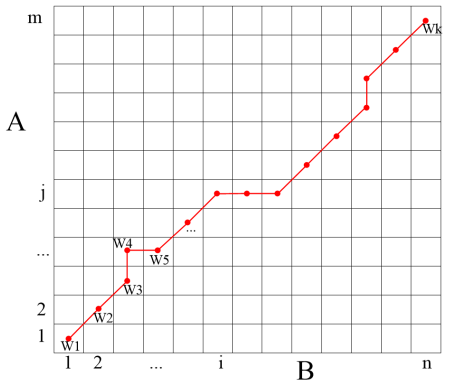

FIGURE 6. Score matrix of dynamic programming. Each point (w_{k+1}) at the warping path (the red) is obtained by the minimum of D(i+1, j), D(i, j+1) and D(i+1, j+1).

b: MODIFICATION OF MAPPING RELATIONSHIP

Suppose two action sequences A and B as defined in Eq. (6), time axis is the X axis, a_i or b_j is the value on Y axis. Traditional DTW algorithm only considers the values on Y axis when mapping the sequences. However, in some case, even $a_i = b_j$, there is no mapping relationship between a_i in A and b_j in B, for a_i is in the upper slope of A and b_j in the lower slope of B. We found that the maximum, minimum and the inflection points of the sequences can be used to improve the dynamic programming process. The new definition of $d(a_i, b_j)$ has incorporated the first and second derivative of the sequence curve.

$$d(a_i,b_j) = w_1 \times (a_i - b_j)^2 + w_2 \times (a_i' - b_j')^2 + w_3 \times (a_i'' - b_j'')^2 \tag{9}$$

where $a_i' = \frac{(a_i - a_{i-1}) + (a_{i+1} - a_{i-1})/2}{2}$ and $a_i'' = a_{i+1} + a_{i-1} - 2a_i$. Parameters w1, w2, w3 must satisfy:

$$w_1 + w_2 + w_3 = 1 \qquad w_1 < w_2, \quad w_1 < w_3$$

c: LIMITATION ON COMPUTATION PATH

The main drawback of DTW is the large computation when the sequences are long. In our experiments, we found for most cases some of the area in the score matrix is not reachable during the DP process. Inspired by the work [35], [36] on improved DTW, we limited the slope of the path within the range from 0.5 to 2, as shown in Fig. 7. In Fig. 7, the slope of line A and B is 2 and slope of line C and D is 0.5. So we have

$$X_a + X_b = N, 2X_a + \frac{1}{2}X_b = M \tag{10}$$

Further, we have the constraints on M and N

$$2M - N \geq 3, \quad 2N - M \geq 2 \tag{11}$$

When it is not satisfied, we think the two sequences are too different to align. In this case, we only need to align X to the $[y_{min}, y_{max}]$ part of Y, where

$$y_{min} = \frac{x}{2}$$

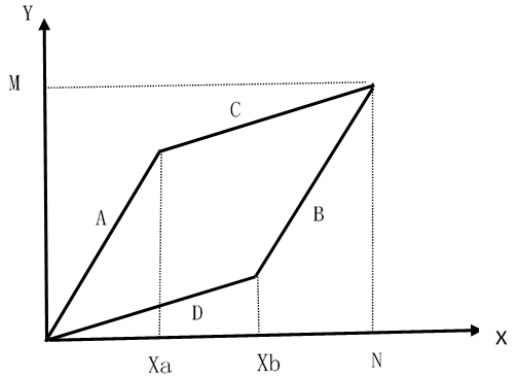

FIGURE 7. Limit path selection in dynamic programming.

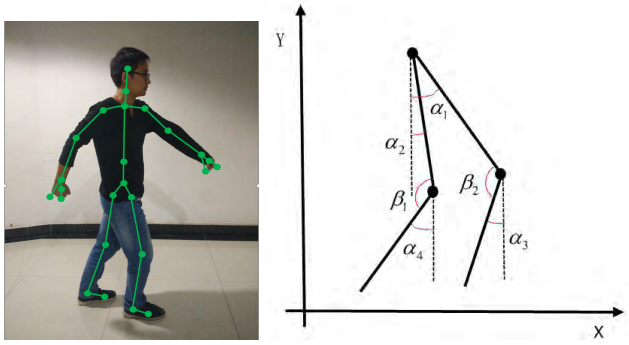

FIGURE 8. Angles for abnormal gait detection.

$$
\begin{aligned}
y_{max} &= 2x \quad x \le X_a \\
y_{min} &= \frac{x}{2} \\
y_{max} &= \frac{x}{2} + (M - \frac{N}{2}) \quad X_a \le x \le X_b \\
y_{min} &= 2x - 2N + M \\
y_{max} &= \frac{x}{2} + (M - \frac{N}{2}) \quad x \ge X_b
\end{aligned} \tag{12}
$$

In this way, the calculation area in the dynamic programming matrix is reduced.

3) ABNORMAL GAIT DETECTION

Abnormal gait detection is to assess if the lower limb movement is abnormal. Because of the occlusion problem during the assessment using Kinect sensors, in our study, we only assessed the gait status of in terms of two states, i.e. normal and abnormal. For the lower limb movement, the hip, knees and ankles are the main participants of the walking process. Therefore, gait features mainly include the motion characteristics such as step length, walking pace, angles of the knee and hip angles. Let X axis be the walking direction, Y axis is vertical downward. We calculated four joint angles $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ related to hip joints and two knee angles β_1, β_2, see Fig. 8. So the feature vector to describe gait is

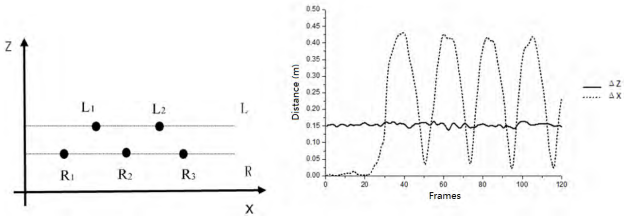

FIGURE 9. Left figure shows the position of the left and right ankle during walking. Right one shows the changes of Δx and Δz. Δx is the ankle position difference in each frame during walking. Δz is the distance between the tester and Kinect.

$[\alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta_1, \beta_2, v, s]$, where v, s are the walking pace and step length, respectively.

Step and pace are calculated according to the relative position of ankles given by the Kinect. As is shown in Fig. 9 (Left), the points correspond to the positions of left and right ankle. A walking process can be expressed as a sequence of position changes in the X-Z plane. Define Δx as the ankle position difference

$$\Delta x_i = L_i - R_i \tag{13}$$

where L, R denote the left and right ankle position, i represents the i-th frame image. Kinect captures many data frames of ankle positions during walking. We calculated a flag for every frame by comparing the two adjacent frames Δx_i and Δx_{i+1}.

$$flag = \begin{cases} 1 & \Delta x_i \neq \Delta x_{i+1} \\ 0 & \Delta x_i = \Delta x_{i+1} \end{cases}$$

We can see that $flag = 1$ means the tester is walking. Each consecutive segment of 1s (flage=1) is a step, and the length of step S_j is maximum of Δx_i within the frames of the step j. The walking pace V could be obtained by

$$V = \frac{n \times \overline{S_j}}{T} \tag{14}$$

where n is the total number of steps and T is the duration of the walking.

Fig. 9 (Right) shows the changes of Δz and Δx, where Δz curve shows the difference in Z direction, which is the distance between the tester and the Kinect. From the curve, we can see it is quite constant, which means the walking direction is parallel to the Kinect X axis. We can see that the step length is about $0.43m$, which is around the peaks of the dashed line.

With the data sampled by Kinect for the lower limb activities, we used MDTW+K-nearest neighbor(KNN) [29] algorithm to compare the sample data of the target to the samples in the database, and found the smallest K samples as the nearest neighbors of the test sample, and then determined the category of the test sample, which is the most frequent category among the K neighbors found from database.

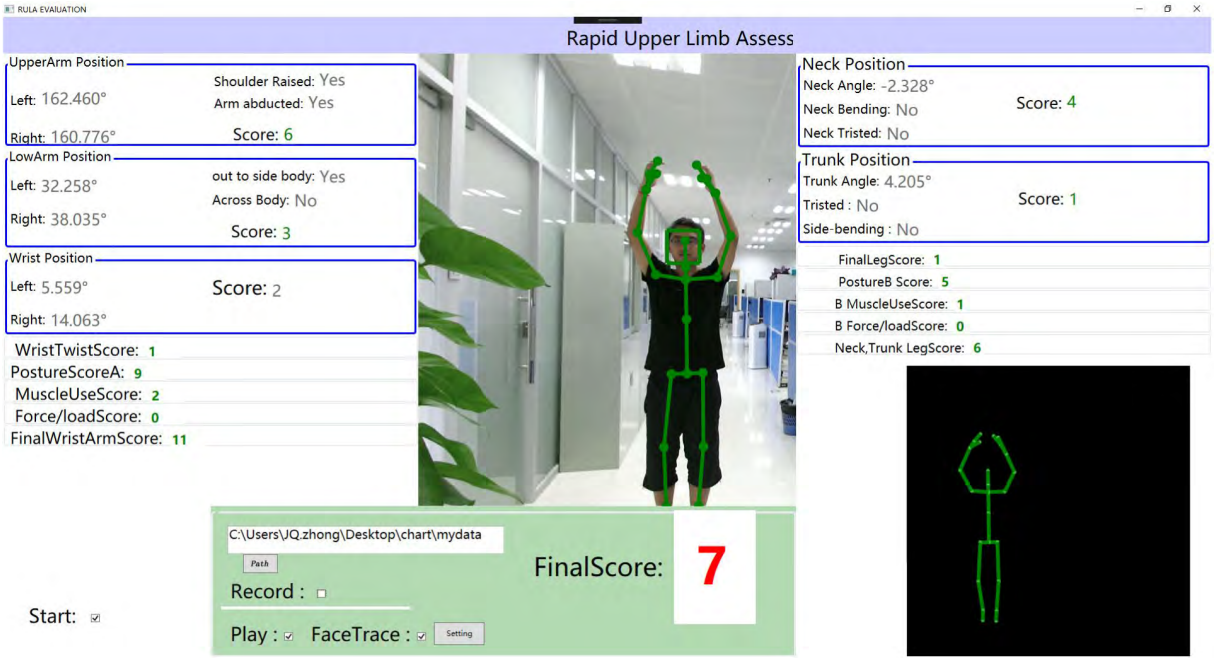

FIGURE 10. Overview of RULA GUI based on Kinect v2. It can automatically capture the assessment parameters and give RULA score in real time. Most of the parameters are from RULA worksheet of Fig. 18. Fig. 11 shows the calculation of neck side bending, trunk side-bending and trunk twist.

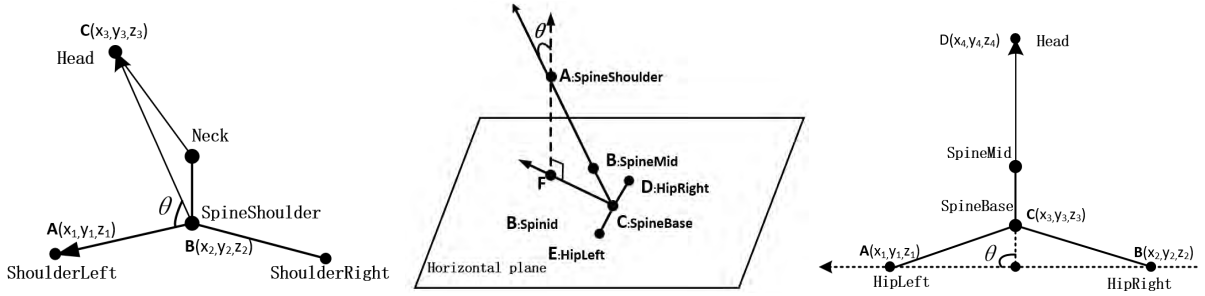

FIGURE 11. Calculation of neck bending (left), trunk angle (middle) and trunk side-bending (right).

C. RULA EVALUATION BASED ON KINECT V2

1) THE RULA METHOD

The RULA is a postural targeting method for estimating the risks of upper limb disorders, which gives a quick and systematic assessment of the postural risks to the tester. The grand score given by RULA based on the movement of those body parts indicates whether or not the intervention is required to reduce the risks of injury. For the simplicity, we do not include the RULA here. Basically, the grand score is calculated using three tables A, B and C. Table A gives the posture score according to the movements of arms and wrists. And table B assesses neck, trunk and legs. Then, Table C takes the scores from A and B as input and calculate the final score. The meaning of the final score is listed as following.

- 1 or 2: the posture is acceptable if it is not maintained or repeated for long periods.
- 3 or 4: the posture should be investigated and making a change.
- 5 or 6: a further investigation is needed and change soon.
- 7: investigate and change immediately.

For more details, please check the RULA worksheet shown in Fig. 18.

2) SEMI-AUTOMATIC RULA BASED ON KINECT V2

Kinect v2 can return a joint skeleton of 25 main joints of human body, as shown in Fig. 2. According to the RULA requirement, 19 of the 25 joints are used to compute RULA paremeters. We designed a software tool based on Kinect for semi-automatic RULA (Fig. 10). The GUI is updated in real-time according to the position of arms, wrists, neck and trunk. For each of these parts, a score is assigned using the RULA tables mentioned above. Then, the final grand score will be obtained and showed in the GUI. The [Record] button allows user to record the whole posture data and save in files for offline analysis. [FaceTrace] button allows user to track

the participant's face, so that we can detect whether the neck is twisted in real time.

Most of the parameters shown in Fig. 10 are directly taken from the RULA worksheet shown in Fig. 18. For Neck section, the angle of neck is defined as the θ in Fig. 4. The calculation of neck bending is shown in Fig. 11. If the angle is within the range of [75°105°], it is false. Neck twist is related to the angle of tuning the head to left or right, if the angle is more than a threshold, usually 5°, it is set to true. For the trunk part, the middle and right sub-figure of Fig. 11 shows the definition of trunk angle and trunk side-bending. If the θ is within the range of [75°, 105°], trunk bending is false. Similar to neck twist, trunk twist is related to the angle of turning the body trunk to left or right. if the angle is more than a threshold, usually 5°, it is set to true.

III. EXPERIMENT AND RESULTS

We evaluated the Kinect v2 based RULA system in the following four directions:

1) evaluate the effectiveness of the joint angles obtained by Kinect by comparing to the digital angle protractor.
2) evaluate the improved motion recognition based on joint angle features.
3) evaluate the performance of gait abnormal detection based on joint angle features.
4) validate the effectiveness of the implemented RULA system using Kinect v2.

All of the experiments were based on the proposed fog framework using Kinect v2. A tower-scale server (T130) with CPU of Intel Core E3-1200V5 3.0 GHz and 4GB of RAM was used as fog server; Microsoft Kinect v2 sensors with a PC (CPU: Intel Core i5-4200 2.50 GHz, 8GB RAM) were used as collecting tools. The Kinect was mounted at 1.2 meters high from and parallel to the ground. The testers are walking parallel to the Kinect camera, and about 3.5 meters away from the camera for RULA test, 2.5 meters for other experiments.

A. COMPARISON WITH DIGITAL ANGLE PROTRACTOR

In order to validate the effectiveness of the joint angles obtained by Kinect, we compared with the measurement using digital angle protractor which is widely used in angle measurement.

20 university students participated in this comparison. The average age was 23.9 ± 1.90; body mass index ranged from 18.2 to 32.8 $kg \cdot m^{-2}$ (mean 24.0 $kg \cdot m^{-2}$, standard deviation 4 $kg \cdot m^{-2}$). None of them was suffering from any physical or mental impairment that would affect this experiment. Written consents were obtained. They were asked to do specific motions in our experiments. The 20 participants were asked to do motions specified in Table 1. While the tester was keeping a pose for a few seconds, we measured the angle for each joint.

The biggest difference is less than 0.76°, the average error is 0.67° and the standard deviation is 0.19°. Table 2 shows the average angle measurement between Kinect and digital angle protractor for each category. From the mean values in the table obtained from the 20 participants, we can see that the two measurements are very close. So, it is good enough to use Kinect sensor to obtain the joint angles.

TABLE 2. Comparison of mean angle measurements between goniometer and our method. The unit is degree.

Degree of Joint Mobility	Goniometer	Ours
Knee extension/flexion	83.63	83.50
Hip extension/flexion	30.98	31.19
Shoulder extension/flexion	119.72	120.10
Elbow extension/flexion	40.22	39.49
Neck extension/flexion	11.62	12.81

B. VALIDATION OF ACTION RECOGNITION

To validate the effectiveness of the modified dynamic time warping (MDTW), we compared the traditional DTW and MDTW in terms of accuracy and efficiency. We tested the methods on two datasets, i.e. SYSU 3D Human-Object Interaction dataset and in-house collected Human Upper Action dataset.

The weight parameters w_1, w_2, w_3 in Eq. (9) were assigned empirically to 0.1, 0.4, 0.5 respectively in the beginning for all the experiments. In order to find which parameter plays a major role in improving action recognition, we changed the proportion of w_1, w_2 and w_3 in the following directions: 1) $w_1 > w_2, w_3$, 2) $w_2 > w_1, w_3$, and 3) $w_3 > w_1, w_2$.

1) TEST ON SYSU 3D HUMAN-OBJECT INTERACTION DATASET

We tested the MDTW method on SYSU 3D Human-Object Interaction Dataset [40], which is commonly used as a standard set for 3D human activity study. It contains 480 video clips of 12 different activities (drinking, pouring, calling phone, etc) performed by 40 individuals. We followed the same experimental setting as other related works. Half of the actions were used as dataset for searching and the other half for testing.

To evaluate our MDTW method, we compared it with the reported results focusing on different feature extraction [37]–[39], including traditional DTW. Table 3 shows the result and comparison. Our method achieves an accuracy of 83.7%, which outperforms other RGB-D activity recognition systems, especially the traditional DTW by a considerable margin (27%), which implies MDTW has improved the dynamic mapping process and reduced the mapping relationship errors effectively.

Weight parameters w_1, w_2, w_3 could change the proportion of the three components in Eq. (9). Table 5 shows how the action recognition accuracy varies with different weight parameters. We can see that the action recognition accuracy reaches the highest accuracy (0.8375) when parameter w_2 plays the major role ($w_2 > w_1; w_3$, i.e. (0.3,0.5,0.2)).

TABLE 3. Accuracy comparison on SYSU 3D Human-Object Interaction set.

	Method	Accuracy(%)
Reported Results	3D Joints and LOP Fourier [37]	78.0
	HON4D [38]	80.0
	SSFF [39]	81.9
Traditional method	Dynamic Time Warping (DTW)	56.7
Ours Results	Modified Dynamic Time Warping (MDTW)	**83.7**

TABLE 4. Accuracy comparison on Human Upper Action dataset.

Method	Accuracy(%)
Dynamic Time Warping (DTW)	80.5
Modified Dynamic Time Warping (MDTW)	89.4

TABLE 5. The recognition accuracy on SYSU 3D Human-Object Interaction dataset with different weights.

w_1	w_2	w_3	Recognition Accuracy(%)	w_1	w_2	w_3	Recognition Accuracy(%)
0.1	0.2	0.7	52.50	0.3	0.5	0.2	**83.75**
0.1	0.4	0.5	35.41	0.355	0.5	0.145	82.35
0.1	0.5	0.4	37.08	0.4	0.5	0.1	81.66
0.1	0.6	0.3	59.58	0.5	0.2	0.3	82.08
0.2	0.3	0.5	40.83	0.5	0.3	0.2	81.66
0.25	0.15	0.6	40.83	0.5	0.4	0.1	43.75
0.3	0.2	0.5	41.25	0.6	0.1	0.3	81.66

TABLE 6. Recognition accuracy of MDTW on Human Upper Action set with different weights.

w_1	w_2	w_3	Recognition Accuracy(%)	w_1	w_2	w_3	Recognition Accuracy(%)
0.05	0.3	0.65	73.25	0.3	0.5	0.2	88.253
0.1	0.3	0.6	81.08	0.4	0.5	0.1	**89.50**
0.1	0.4	0.5	81.16	0.5	0.2	0.3	83.30
0.2	0.3	0.5	84.33	0.5	0.3	0.2	80.58
0.3	0.2	0.5	82.75	0.5	0.4	0.1	83.30
0.3	0.4	0.3	86.58	0.5	0.45	0.05	88.74
0.3	0.45	0.25	87.66	0.6	0.3	0.1	88.25

The confusion matrix of the results by our MDTW is presented in Fig. 12. It can be seen that MDTW often confuses the activity of sweeping with mopping, as the two are similar. In addition, taking from wallet occasionally is also misidentified as calling phone and playing phone due to the similarity among the three.

2) TEST ON HUMAN UPPER ACTION DATASET

We collected a new dataset containing upper body actions using Microsoft Kinect v2 camera to evaluate MDTW. We named this as Human Upper Action dataset (HUA). To collect this set, 10 subjects were asked to perform 12 different actions, three times for each. The average age was 22.7±1.30; body mass index ranged from 19.3 to 28.3 $kg{\cdot}m^{-2}$ (mean 23.8 $kg \cdot m^{-2}$, standard deviation 3.9 $kg \cdot m^{-2}$) The 12 actions are cross-waving, raising hand, shoulder turning etc. as shown in Fig. 13. So, this dataset has in total 360 skeleton clips. For the test, for each action, half of the samples are used for testing and the rest form the dataset for searching.

The comparison result is shown in Table 4. MDTW method achieves the accuracy of 89.5%, significantly better than that of the traditional DTW by a large margin (9.0%). Changing the values of w_1, w_2, w_3, we can see from Table 6 that the action recognition accuracy reaches the highest (0.895) when $w_1, w_2, w_3 = 0.5, 0.4, 0.1$, (the first and second order differential play the major part), which implies the fusion of original values, first and second order differential with a particular weight can improve the dynamic matching process and recognition accuracy.

By examining the confusion matrix of MDTW in Fig. 14, we observed that MDTW method can distinguish most of the actions well, including walking, arm swing, TW stretching (stretching out hands to resemble 'T' or 'W'), shoulder shaking and cross-waving. However, we noticed that the action of striking is often confused with TW stretching by MDTW. This is because the action of TW stretching is quite similar with the action of striking. The same for swing up and down.

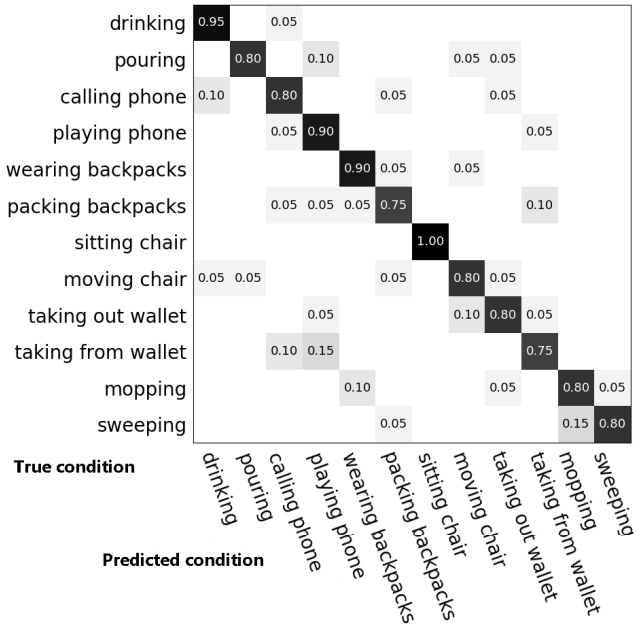

FIGURE 12. Confusion matrix of test MDTW on SYSU 3D Human-Object Interaction dataset.

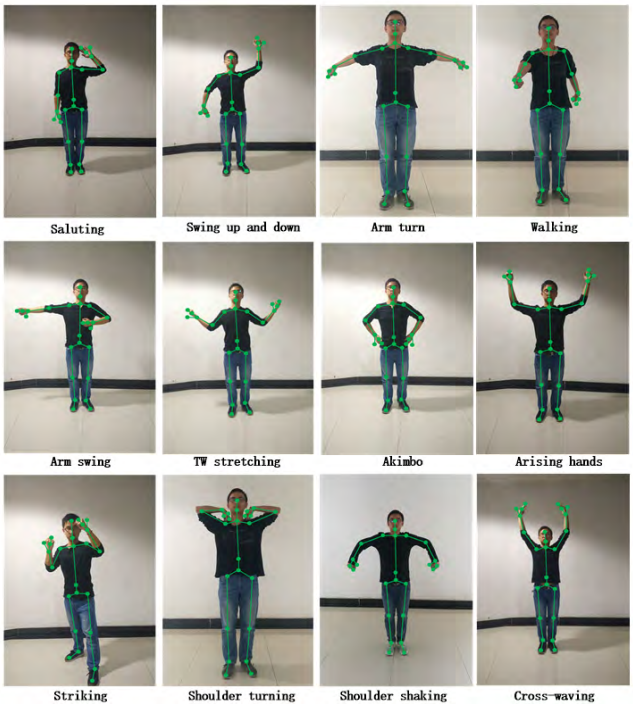

FIGURE 13. Snapshots of twelve actions in HUA dataset. The skeleton data are indicated in RGB channels.

3) COMPARISON OF COMPUTATION TIME

In MDTW method, we limited the search space during the dynamic programming process to speed up the running time. Here we show more detailed comparison on the computation complexity of the method, which is mainly the calculating time to obtain the score matrix. Suppose an action sequence

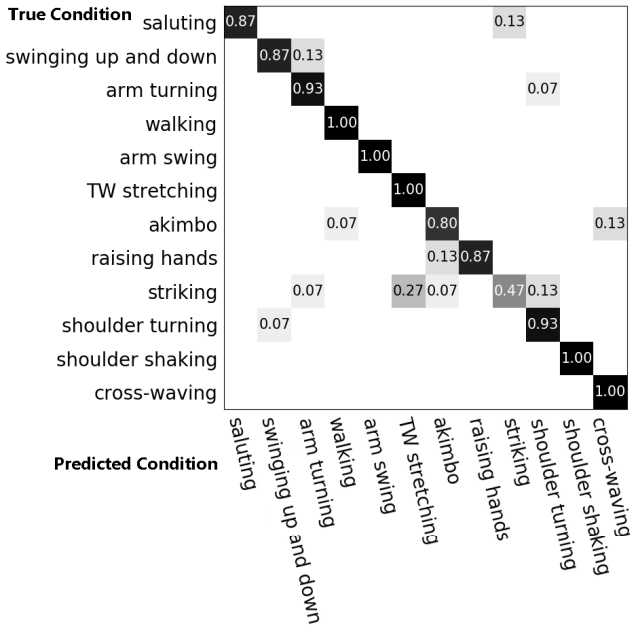

FIGURE 14. Confusion matrix of test MDTW on Human Upper Action dataset.

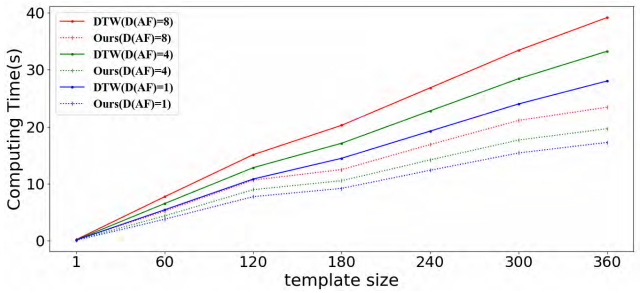

FIGURE 15. Comparison of computation time in seconds between MDTW and DTW with different feature size. The X-axis is the template size, Y-axis is the computation time in unit of second.

contains 80 frames, each with 8 action features. We search the action sequence against the HUA dataset, which makes alignments between the query action sequence and database action sequences. Fig. 15 shows how the computation complexity changes with the dataset size and action feature size (AF = 1, 4, 8). We can see that the computation time reduces using MDTW method. Especially, with the increasing size of the dataset, the gap between the traditional DTW and MDTW increases significantly.

C. EVALUATION OF GAIT ABNORMALITY DETECTION

18 subjects participated in this experiment. The average age was 22.8 ± 1.20, body mass index ranged from 19.3 to 31.1 $kg \cdot m^{-2}$ (mean 23.1 $kg \cdot m^{-2}$, standard deviation 3.92 $kg \cdot m^{-2}$). The data collected from 10 of them was used as the database, the remaining 8 people for testing. Each participant was required to walk 20 times in the direction parallel to Kinect. Among the 20 walkings, 10 were normal and the other

TABLE 7. Simulated anomaly walking. If the step length or duration of left and right leg are significantly different, it is treated as anomaly.

	Short Step Length	Short Step Time
$left \gg right$	abnormal	abnormal
$left \ll right$	abnormal	abnormal
$left \simeq right$	normal	normal

TABLE 8. Results of gait anomaly detection.

	Normal Walking	Gait Anomalies
Correct detections	75	74
Accuracy	93.13%	
Sensitivity	93.75%	
Specificity	92.50%	
Precision	92.59%	

10 were simulated gait anomalies. Therefore, in the database, we have 200 gait sequences, 100-150 frames for each gait sequence. 10-fold cross-validation was used to find out the optimal value of K for KNN algorithm, which is a number from 1 to 9.

Many parameters can be used to describe human gait from different perspectives [41]. We considered two meaningful parameters, short step length and short step time, to control the abnormal gait simulation. Short step length refers to the distance between two successive placements of the same foot while short step time is the time spending on one step with the same foot. According to Table 7, we simulated the abnormal gait in two ways with respect to the two control parameters: 1) taking bigger steps with one of the legs(a empirical threshold set to about 15cm); 2) spend more time on steps of one leg(a empirical threshold set to about 3s).

Table 8 shows the gait detection results from the 8 testers, i.e. 80 sequences for normal walking and 80 for anomalies. The table shows the detection accuracy with different metrics. It can be observed that the detection is very accurate whose mean accuracy is 93.13% and the sensitivity is 93.75%.

D. EVALUATION OF RULA USING KINECT V2 IN PFA

Due to the strong motion capture capability and the affordable cost of Kinect v2, it has been proved to be a good choice for RULA by several works that conducted semi-automatic RULA [27], [42], [43]. Therefore, we just do a brief validation for RULA with Kinect in this part. One major issue of this scheme is that one has to estimate the neck twist manually during the RULA process, which introduces a significant basis into the evaluation. In our study, the joint orientation data provided by Kinect was used to calculate the status of the face and the neck. Usually, if the neck's twist exceeds a predefined threshold (which usually set to 5°), we say the neck is twisted. This makes the RULA evaluation more accurate.

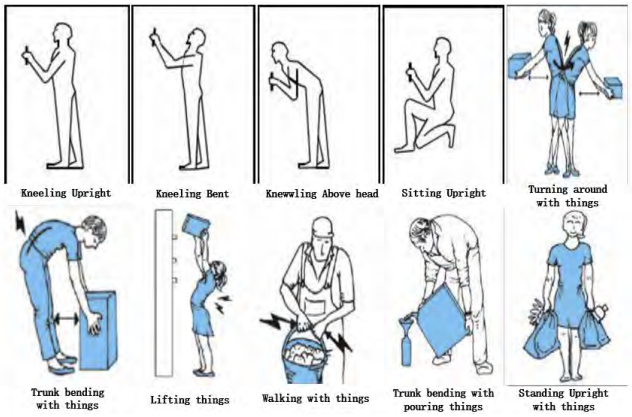

FIGURE 16. Ten actions numbered sequentially from left to right.

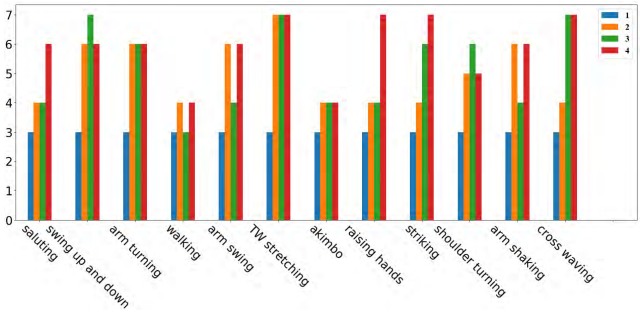

FIGURE 17. RULA score of action state for 12 actions from HUA dataset, Fig. 13.

1) VALIDATION OF THE RULA USING KINECT V2

We compared this semi-automatic RULA system with an expert to validate its effectiveness in fog computing framework. To do this, we selected 10 postures as shown in Fig. 16, which are related to musculoskeletal disorders in workplace [27]. These 10 postures were performed by three participants, whose average age was 22.3 $\pm$ 0.47 and body mass index are 19.3,21.9 and 26.3 respectively. We also invited one RULA expert, who is a professional doctor with more than 10 years of practice. More experts may produce statistically stabler result, but Sara et al. conducted an investigation of RULA reliability and demonstrated that intra-rater reliability is higher than the inter-rater reliability [44], which implies that serial assessments would be more consistent if carried out by the same person.

During the processing, the software tool kept tracking the tester's face and calculating the RULA parameters and the RULA final score of the tester's posture in real time. The participants did the tests 10 times. Table 9 shows the rounded average RULA grand score by the software compared with the scores by the RULA expert. We can see that the two sets of scores shows a good agreement, especially for action one, four, six, seven, eight and ten. In action two, expert's score is lower than Kinect score because the expert cannot get the position of the neck precisely, which leads to the underestimation of the ergonomic risks. In action three, five and nine, There are some differences between the Kinect score and

TABLE 9. Comparison between RULA software and expert evaluation. The numbers are rounded average scores of the tests repeated 10 times. Lower score means better health conditions.

	Action1	Action2	Action3	Action4	Action5	Action6	Action7	Action8	Action9	Action10
Expert	3	4	6	4	4	7	7	4	6	2
Tester-1	3	6	5	4	6	7	7	4	5	2
Expert	2	4	6	4	3	7	7	4	5	2
Tester-2	2	7	5	4	5	7	7	4	5	2
Expert	3	3	5	4	4	7	7	4	5	2
Tester-3	3	7	4	4	5	7	7	4	4	2

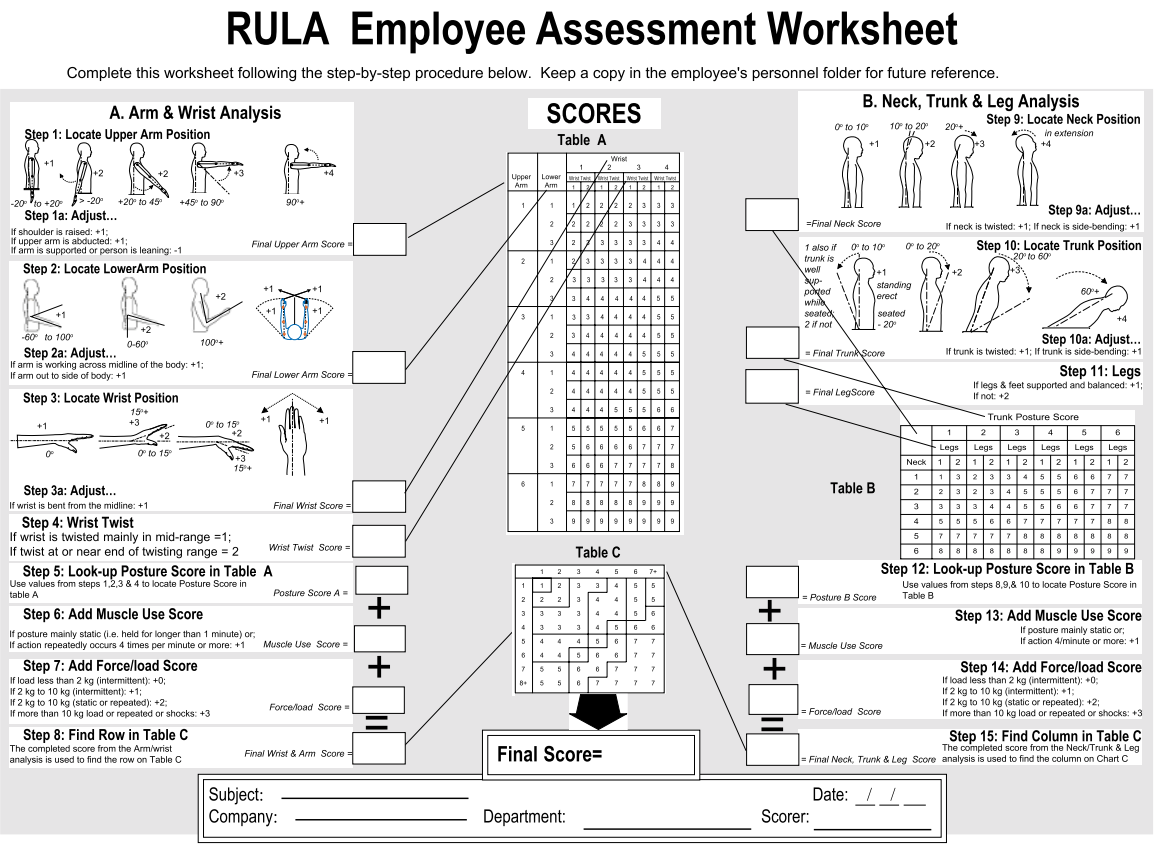

FIGURE 18. RULA worksheet.

expert because of the occlusion during the motion capture. On average, the grand-scores given by the tool agree with that of the expert very well, except for the first a few actions in Table 9. In terms of the difference between Kinect tool and the expert, for tester 1, the mean difference is 0.2, standard deviation (std) is 0.98; for tester 2, the mean difference is 0.3, std is 1.345; for tester 3, the mean is 0.4 and std is 1.113.

Except for the occlusion problem, the modified semi-automatic RULA system can be used effectively compared with a RULA expert evaluation and reduce the bias from the expert for some actions.

2) TEST SEMI-AUTOMATIC RULA IN PFA FRAMEWORK

We used the semi-automatic RULA system to make evaluation for the human body upper limb motions. The actions in HUA dataset was used for the experiment. In this test, we did not consider any load for arms and legs, i.e. choose the option of less than 2kg in RULA. Each tester started with hands down as the initial state. During each of the actions, we recorded the changes in the RULA score. For each action in HUA dataset (Fig. 13), we split the period into four roughly equal segments. Each segment corresponds to a state or moment.

Fig. 17 shows the RULA scores for each state of the twelve actions. We can see that RULA scores are relatively low for the beginning states for all twelve actions. These actions have different RULA scores, which means different level of ergonomic risks. For example, activities of walking and akimbo have lower grand-score than strenuous actions such as arm swing, arm up, striking and arm cross waving. And strenuous exercise have higher risks.

IV. CONCLUSIONS

In this paper, we presented a framework based on fog computing and described how it works for families. Under this framework, we first evaluated the accuracy for joints angles using Kinect, compared with digital angle protractor. Then we modified the traditional DTW algorithm improving the measurements of the similarity between two actions and accelerating the calculation. Tests of motion recognition on two datasets(one is collected by autors) showed a better recognition accuracy. To test the gait abnormal detection, we set up a gait dataset using RULA and showed high detection accuracy with the modified DTW algorithm and simple KNN method. Finally, we evaluated the RULA software tool to show that the high agreement between the automatic RULA and RULA by human expert. The overall results showed the effectiveness of physiological function assessment using Kinect v2 in fog computing environment.

To extend the work in the future, we are going to deploy the fog servers in medical centers, collect real data from elderly to validate the effectiveness of proposed physiological function assessment framework and the performance of the function components such as motion recognition and gait abnormal detection. Another future work would be to enrich the methods assess the physiological conditions of the elderly, for example, to measure TUG (Time Up and Go) to assess their walking.

ACKNOWLEDGEMENT

The authors thank the ICME 2018 Workshop Chair, Abdur Rahman, for the constructive advice.

REFERENCES

[1] M. S. Hossain and G. Muhammad, "Cloud-assisted industrial Internet of Things (IIoT)—Enabled framework for health monitoring," *Comput. Netw.*, vol. 101, pp. 192–202, Jun. 2016.
[2] L. Zhu, L. Li, and C. Meng, "Construction and analysis of a monitoring system with remote real-time multiple physiological parameters based on cloud computing," *J. Biomed. Eng.*, vol. 31, no. 6, pp. 1377–1383, 2014.
[3] *Gartner Says the Internet of Things Will Transform the Data Center[EB/OL]*. Accessed: Jun. 28, 2016. [Online]. Available: http://www.gartner.com/newsroom/id/2684616
[4] S. Mehraghdam and H. Karl, "Specification of complex structures in distributed service function chaining using a YANG data model," 2015, *arXiv:1503.02442*. [Online]. Available: https://arxiv.org/abs/1503.02442
[5] M. T. Nagy, R. J. MacFarlane, Y. Khan, and M. Waseem, "The frozen shoulder: Myths and realities," *Open Orthopaedics*, vol. 7, pp. 352–355, Sep. 2013.
[6] E. Weiss and R. Jurmain, "Osteoarthritis revisited: A contemporary review of aetiology," *Int. J. Osteoarchaeol.*, vol. 17, no. 5, pp. 437–450, 2010.
[7] J. A. Kanis, L. J. Melton, III, C. Christiansen, C. C. Johnston, and N. Khaltaev, "The diagnosis of osteoporosis," *J. Bone Mineral Res.*, vol. 9, no. 8, pp. 1137–1141, 1994.
[8] N. E. Mayo, S. Wood-Dauphinee, S. Ahmed, G. Carron, J. Higgins, S. Mcewen, and N. Salbach, "Disablement following stroke," *Disab. Rehabil.*, vol. 21, nos. 5–6, pp. 258–268, 1999.
[9] C. Hu, Y. Ge, and Y. Chen, "Human joint mobility measurement system," *Chin. J. Med. Phys.*, vol. 33, no. 1, pp. 34–38, 2016.
[10] M. Parajuli, D. Tran, W. Ma, and D. Sharma, "Senior health monitoring using Kinect," in *Proc. 4th Int. Conf. Commun. Electron.*, Aug. 2012, pp. 309–312.
[11] R. Z.-L. Hu, A. Hartfiel, J. Tung, A. Fakih, J. Hoey, and P. Poupart, "3D pose tracking of walker users' lower limb with a structured-light camera on a moving platform," in *Proc. CVPR Workshops*, Jun. 2011, pp. 29–36.
[12] P. J. Standen, D. J. Brown, S. Battersby, M. Walker, L. Connell, A. Richardson, F. Platts, K. Threapleton, and A. Burton, "A study to evaluate a low cost virtual reality system for home based rehabilitation of the upper limb following stroke," *Int. J. Disab. Hum. Develop.*, vol. 10, no. 4, pp. 337–341, Nov. 2010.
[13] S. Li and B. Xu, "Application of Kinect camera in rapid upper limb evaluation," *Machinery*, vol. 55, no. 10, pp. 36–40, 2017.
[14] C.-C. Chen, "Improvement in the physiological function and standing stability based on Kinect multimedia for older people," *J. Phys. Therapy Sci.*, vol. 28, pp. 1343–1348, Apr. 2016.
[15] A. Bernardino, C. Vismara, S. B. I Badia, E. Gouveia, F. Baptista, F. Carnide, S. Oom, and H. Gamboa, "A dataset for the automatic assessment of functional senior fitness tests using Kinect and physiological sensors," in *Proc. 1st Int. Conf. Technol. Innov. Sports, Health Wellbeing (TISHW)*, Dec. 2016, pp. 1–6.
[16] H. Sakoe and S. Chiba, "A similarity evaluation of speech patterns by dynamic programming," in *Proc. Nat. Meeting Inst. Electron. Commun. Eng. Jpn.*, Jul. 1970, p. 136.
[17] H. Sakoe and S. Chiba, "A dynamic programming approach to continuous speech recognition," in *Proc. Int. Congr. Acoust.*, Budapest, Hungary, vol. C-13, Jul. 1971, pp. 139–142.
[18] H. Sakoe and S. Chiba,, "Comparative study of DP-pattern matching techniques for speech recognition," in *Proc. Tech. Group Meeting Speech, Acoustic Soc. Jpn.*, 1973, pp. 450–455.
[19] M. Müller, "Dynamic time warping," *Inf. Retr. Music Motion*, vol. 52, pp. 69–84, 2007.
[20] L. McAtamney and E. N. Corlett, "RULA: A survey method for the investigation of work-related upper limb disorders," *Appl. Ergonom.*, vol. 24, no. 2, pp. 91–99, 1993.
[21] D. Roman-Liu, "Comparison of concepts in easy-to-use methods for MSD risk assessment," *Appl. Ergonom.*, vol. 45, no. 3, pp. 420–427, 2014.
[22] R. A. Clark, Y.-H. Pua, K. Fortin, C. Ritchie, K. E. Webster, L. Denehy, and A. L. Bryant, "Validity of the microsoft Kinect for assessment of postural control," *Gait Posture*, vol. 36, no. 3, pp. 372–377, 2012.
[23] R. A. Clark, K. J. Bower, B. F. Mentiplay, K. Paterson, and Y.-H. Pua, "Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables," *J. Biomech.*, vol. 46, no. 15, pp. 2722–2725, Aug. 2013.
[24] T. Dutta, "Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace," *Appl. Ergonom.*, vol. 43, no. 4, pp. 645–649, 2012.
[25] B. Bonnechère, B. Jansen, P. Salvia, H. Bouzahouene, L. Omelina, F. Moiseev, V. Sholukha, J. Cornelis, M. Rooze, and S. S. J. Van, "Validity and reliability of the Kinect within functional assessment activities: Comparison with standard stereophotogrammetry," *Gait Posture*, vol. 39, no. 1, pp. 593–598, Jul. 2014.
[26] H. Haggag, M. Hossny, S. Nahavandi, and D. Creighton, "Real time ergonomic assessment for assembly operations using Kinect," in *Proc. UKSim 15th Int. Conf. Comput. Modelling Simulation*, Apr. 2013, pp. 495–500.
[27] V. M. Manghisi, A. E. Uva, M. Fiorentino, V. Bevilacqua, G. F. Trotta, and G. Monno, "Real time RULA assessment using Kinect v2 sensor," *Appl. Ergonom.*, vol. 65, pp. 481–491, Nov. 2017.
[28] L. E. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, p. 1883, 2009.
[29] Y. Guo, Y. Li, and Z. Shao, "DSRF: A flexible trajectory descriptor for articulated human action recognition," *Pattern Recognit.*, vol. 76, pp. 137–148, Apr. 2018.
[30] W. Cao, J. Zhong, G. Cao, and Z. He, "Physiological function assessment based on RGB-D camera," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2018, pp. 1–6.

[31] M. Naeem, R. K. Khushnood, and S. Arshad, "A challenging demand side power management through smart embedded meter system," *Int. J. Comput. Appl.*, vol. 106, no. 17, pp. 7–13, 2014.
[32] N. M. DiFilippo and M. K. Jouaneh, "Characterization of different microsoft Kinect sensor models," *IEEE Sensors*, vol. 15, no. 8, pp. 4554–4564, Aug. 2015.
[33] T.-L. Le, M.-Q. Nguyen, and T.-T.-M. Nguyen, "Human posture recognition using human skeleton provided by Kinect," in *Proc. Int. Conf. Comput., Manage. Telecommun.*, Jan. 2013, pp. 340–345.
[34] H. Rahmani, A. Mahmood, D. Q. Huynh, and A. Mian, "Real time action recognition using histograms of depth gradients and random decision forests," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Mar. 2014, pp. 626–633.
[35] J. Hu, R. Chen, and Z. Li, "Discussion of improved DTW algorithm in speed recognition," *Inf. Technol. Netw. Secur.*, vol. 30, no. 3, pp. 30–32, 2011.
[36] J. Liu and S. Hu, "Discrimination of ST segments in ECG signal based on improved dynamic time warping," *Beijing Biomed. Eng.*, vol. 29, no. 5, pp. 492–496, 2010.
[37] J. Wang, Z. Liu, and Y. Wu, *Learning Actionlet Ensemble for 3D Human Action Recognition*. Springer, 2014.
[38] O. Oreifej and Z. Liu, "HON4D: Histogram of oriented 4D normals for activity recognition from depth sequences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 716–723.
[39] A. Shahroudy, G. Wang, and T.-T. Ng, "Multi-modal feature fusion for action recognition in RGB-D sequences," in *Proc. 6th Int. Symp. Commun., Control Signal Process.*, May 2014, pp. 1–4.
[40] J.-F. Hu, W.-S. Zheng, J. Lai, and J. Zhang, "Jointly learning heterogeneous features for RGB-D activity recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 11, pp. 2186–2200, Nov. 2017.
[41] A. Muro-De-La-Herran, B. Garcia-Zapirain, and A. Mendez-Zorrilla, "Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications," *Sensors*, vol. 14, no. 2, pp. 3362–3394, 2014.
[42] P. Poli, S. B. Healy, and D. P. Dee, "Pose estimation with a Kinect for ergonomic studies: Evaluation of the accuracy using a virtual mannequin," *Sensors*, vol. 15, no. 1, pp. 1785–1803, 2015.
[43] P. Plantard, H. P. H. Shum, A.-S. Le Pierres, and F. Multon, "Validation of an ergonomic assessment method using Kinect data in real workplace conditions," *Appl. Ergonom.*, vol. 65, pp. 562–569, Nov. 2017.
[44] S. Dockrell, E. O'Grady, K. Bennett, C. Mullarkey, R. McConnell, R. Ruddy, S. Twomey, and C. Flannery, "An investigation of the reliability of rapid upper limb assessment (RULA) as a method of assessment of children's computing posture," *Appl. Ergonom.*, vol. 43, no. 3, pp. 632–636, 2012.

WENMING CAO received the M.S. degree from the System Science Institute, China Science Academy, Beijing, China, in 1991, and the Ph.D. degree from the School of Automation, Southeast University, Nanjing, China, in 2003. From 2005 to 2007, he was a Postdoctoral Researcher with the Institute of Semiconductors, Chinese Academy of Sciences. He is currently a Professor with Shenzhen University, Shenzhen, China. He has authored or co-authored over 80 publications in top-tier conferences and journals. His research interests include pattern recognition, image processing, and visual tracking.

JIANQI ZHONG is currently pursuing the M.Eng. degree in integrated circuit engineering from Shenzhen University, Shenzhen, China. His research interests include pattern recognition and physiological function assessment.

GUITAO CAO received the M.S. degree from Shandong University, Jinan, China, in 2001, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2006. She is currently an Associate Professor with the School of Computer Science and Software Engineering, East China Normal University, Shanghai, China. She has authored over 30 publications, including the IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING. Her research interests include image processing and pattern recognition, media analysis, and understanding.

ZHIQUAN HE received the M.S. degree from the Institute of Electronics, Chinese Academy of Sciences, in 2001, and the Ph.D. degree from the Department of Computer Science, University of Missouri, Columbia, MO, USA, in 2014. He is currently an Assistant Professor with the College of Information Engineering, Shenzhen University, China. His research interests include the areas of image processing, computer vision, and machine learning.