A Human Skeleton Data Optimization Algorithm For Multi-Kinect

Hanzhen Zhang
Institute of Artificial Intelligence and Robotics
Xi'an Jiaotong University
Xi'an, China
xcxzhz0152@stu.xjtu.edu.cn

Xiaojuan He
School of Software Engineering
Xi'an Jiaotong University
Xi'an, China
hexiaojuan@stu.xjtu.edu.cn

Yuehu Liu
Institute of Artificial Intelligence and Robotics
Xi'an Jiaotong University
Xi'an, China
liuyh@xjtu.edu.cn

***Abstract*—As a motion capture device, Kinect can extract human skeleton data (three dimensional positions of body joints) through depth information, which has been increasingly applied in many fields such as action recognition, medical diagnosis and rehabilitation evaluation. But when collecting human skeleton data with a single Kinect, the detection range is limited and stale data exceptions or data missing might occur due to the influence of human body occlusion and environment. These invalid data will directly affect the accuracy of relevant calculation results and hinder the application of Kinect. To solve the above problems, we propose a data acquisition scheme based on multi-Kinect and an optimization algorithm for human skeleton data. First, we use the method based spatial region constraints and K-means clustering to remove the outliers in the human skeleton data collected by each Kinect. Then we convert the data from different perspectives into the same coordinate system and perform adaptive weighted fusion. We finally obtain a stable and valid data sequence, facilitating its application greatly. We design relevant experiments to compare the optimized data with the original data and the results verify the effectiveness of the algorithm.**

Keywords—multi-kinect, skeleton data, outlier detection, data fusion

I. Introduction

The change of human body's posture can be described by the trajectory of the joints. The 3D human skeleton model can be constructed by extracting human skeleton data (3D positions of joints) from image information. It can intuitively reproduce the process of human body movement, which can be applied in the fields of gait analysis [1], action recognition [2] and has been widely used in athlete posture correction and clinical medicine rehabilitation training [3,4]. Depth image acquisition technology has developed rapidly in the past few years. Kinect, released by Microsoft, is a new device that can synchronously acquire RGB images and depth information. Compared with VICON and other sports data acquisition systems, Kinect has the advantages of low price, simple operation and strong applicability. Kinect's human joints detection algorithm is encapsulated in the SDK by researchers from Microsoft company who use random forest algorithm to train mass data, which can acquire the spatial 3D coordinates of 25 human joints in real time. The validity of the algorithm and the accuracy of the detection results have been proved in the relevant articles [5,6]. The human skeleton model constructed by such spatial joint points has no background interference and simple features. Therefore, the research and application of posture recognition and analysis algorithm based on Kinect's human skeleton data show an increasing trend year by year.

However, Kinect has some inherent limitations. The measurement range of one Kinect is limited and Kinect will be affected by human body occlusion and environment in practical applications, resulting in abnormal and wrong data of 25 human joints, which seriously affects the accuracy and reliability of motion analysis and other calculations. To overcome the limitations mentioned above, we propose a data acquisition scheme based on multiple Kinects for data optimization and fusion, taking advantage of the redundancy of multi-view data. Firstly, we detect and remove the outliers in the human skeleton data, which is the basic of data fusion. Then we use an adaptive weighted fusion algorithm to fuse each set of skeleton data collected by different Kinect sensors into a stable and valid data sequence to reduce the influence of human body occlusion and improve the robustness and stability of skeleton data. The weight of the Kinect is dynamic, which means Kinect can adjust the weight based on the measured values at different times. This data fusion algorithm can make the fusion results have linear unbiased minimum variance.

To summarize, our main contributions are two-fold:

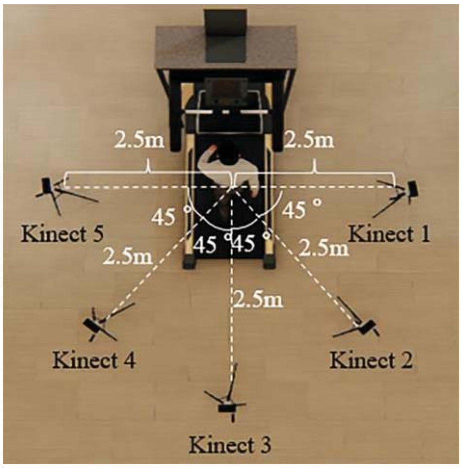

Fig. 1. Sketch of data acquisition scheme.

(1) We adopt a multi-Kinect collection scheme to obtain multi-view human skeleton data to overcome the limitations of a single Kinect.

(2) We propose a multi-view Kinect skeleton data

optimization algorithm based on K-means clustering and adaptive weighted fusion. And experimental results verify the validity of the algorithm.

II. PRBLEM FORMULATION

A single Kinect V2 has a limited tracking range for the human body. A Kinect provides a limited field of view (FOV) of 43° in vertical and 57° in horizontal, while the effective range in z axis is 1 to 3 meters [7]. The accuracy of the human skeleton data will be affected if the subject is too close or too far away from it. If subjects exceeds this range, the noise will become larger as the distance increases, which will result in inaccurate or even seriously abnormal detection of human joints. To overcome the limitations mentioned above, researchers have taken different approaches using multiple Kinect. e.g., in [8], the authors show that under ideal conditions using several Kinect sensors increases the precision of the data collected. In another attempt [9], a series of Kinect sensors were placed at a certain distance on only one side of the walking path. In [10], subjects were asked to walk on a treadmill in order to keep their distance to the Kinect within effective range at all times. Considering the skeleton data collected by a single Kinect may be abnormal or missing, we propose a multi-Kinect data collection scheme based on treadmill to improve the robustness and effectiveness of the data and obtain a complete and continuous human movement data sequence. As shown in Fig. 1, five Kinect sensors were placed in different positions on the subject's side and behind (More Kinect can be used if the actual environment allows). They are 2 to 2.5 meters away from the subjects, ensuring that the subject is within the effective detection range of Kinect. The subject is asked to walk on a treadmill and five Kinect collected human skeleton data at the same time from different perspectives, which could not only obtain the coordinate sequence of body joints, but also overcome the problem that partial data collected by a single Kinect is invalid.

Then, how to deal with the three-dimensional skeleton data of human body from multiple perspectives and obtain a stable and effective data sequence of human joints is the main problem to be solved in this paper. It main include two aspects: (1) Kinect will obtain abnormal and wrong skeleton data due to the influence of human body occlusion and measuring environment. These invalid data will produce error transmission and directly affect the result of coordinate transformation and data fusion. Therefore, it is particularly necessary to detect outliers before data fusion. (2) Kinect's algorithm for detecting human body joints is a trained estimation algorithm essentially, which has inherent measurement error. To increase the robustness and overall validity of the data and obtain a complete and continuous data sequence, we need to fuse the data obtained by different Kinect. There is a lot of literature on outliers detection which describes a variety of approaches. e.g., the method based LOF [11]. This method is difficult to select parameters and the calculation is complex. In [12], the author proposes a distance-based outlier detection method. It is not suitable if the data set is large or has different regional densities. Some outlier detection methods based clustering are proposed in [13,14]. Nevertheless, there are not many researches on Kinect skeleton data fusion algorithm. In [15], the authors used two Kinect to track the subjects' body joints and kalman filter algorithms is used to process skeleton data. However, the noise characteristics of skeleton data are not analyzed in this paper, and there is no theoretical basis for using the kalman filter to process skeleton data directly. In [16], the author proposed a data fusion algorithm based extended set membership filter (ESMF), but the observation equation should be known. All of the methods mentioned above require prior knowledge. In reality, we can't or hardly get the corresponding prior knowledge of Kinect sensors. This is a huge difficulty for the processing of skeleton data and the application of Kinect. There have been many researches and applications about multi-sensor data fusion [17,18], which can be applied to the fusion of Kinect skeleton data. In [17], the authors proposed a multi-sensor weighted fusion algorithm based on unbiased estimation.

To solve the problems mentioned above, this paper proposes an outlier detection algorithm based on spatial region constraints and K-means clustering and adopts an adaptive weighted fusion algorithm to fuse the human skeleton data collected by multiple Kinect without any prior knowledge of measurement data.

Algorithm 1 The procedure of outliers detection for the coordinates sequence of each joints.

Initialize: $\mathbf{p}=(x_p, y_p, z_p)$ denote the 3D coordinates of point p, S denote the set of points within a constrained space, N is the number of input points, K is the number of clusters, n_i is the number of points contained in the cluster-i, $i \in \{1,2,\ldots,K\}$, p_c is the center, δ is the threshold value

1: Define the constrained space of the joint

2: **if** $p \notin S$

  $\mathbf{p}=\mathbf{0}$ (This point is identified as an outlier.)
 end if

3: Perform K-means clustering for the remaining points.

4: **while** $p \in$ cluster-i

 if $n_i \le \dfrac{N}{3K}$

  $\mathbf{p}=\mathbf{0}$

 else

 Compute $d_j=\|\mathbf{p}-\mathbf{p_c}\|$ ($j \in \{1,2,\ldots, n_i\}$), $\mu=\frac{1}{n_i}\sum_{j=1}^{n_i} d_j$,

$\sigma=\sqrt{\frac{1}{n_i}\sum_{j=1}^{n_i}(x_i-\mu)^2}$, $\delta=\mu+3\sigma$

 if $d_j \ge \delta$

  $\mathbf{p}=\mathbf{0}$

 end if

 end if

 end while

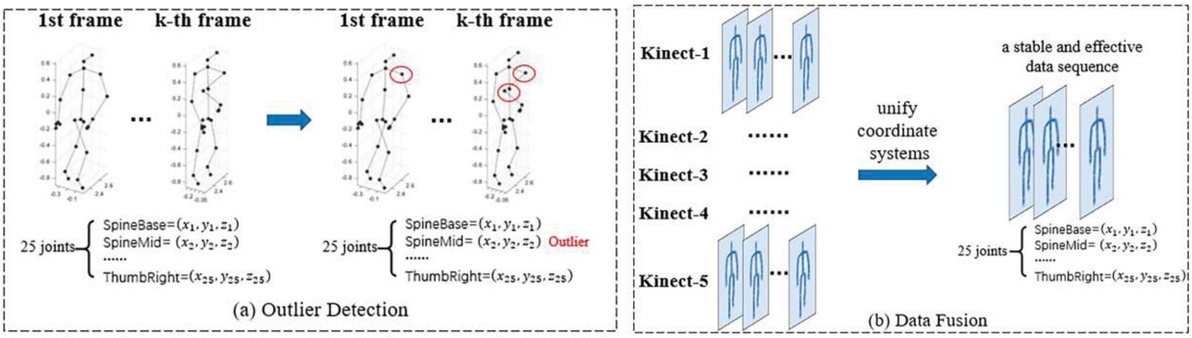

Fig. 2. Overview of framework.

III. Method

A. Overview of Algorithm Framework

As shown in Fig. 2, we first obtain the coordinate sequences of subject's joints from multi perspectives with five Kinect sensors and computers. Then we use the method based K-means clustering to detect and remove the outliers in the human skeleton data collected by each Kinect and convert the 3D coordinates of body joints from five perspectives into Kinect3's coordinate system. At last, we perform adaptive weighted fusion to obtain a stable and effective data sequence.

B. Outliers Detection for Skeleton Data

Considering the characteristics of the skeleton data collected by Kinect in the actual environment, this paper proposes an outlier detection algorithm based on spatial region constraints and K-means clustering. Clustering-based methods detect outliers by examining the relationship between objects and clusters. Intuitively, an outlier is an object that belongs to a small and remote cluster, or does not belong to any cluster. The method can be summarized as the pseudo-code in Algorithm. 1. Combined with the movement characteristics of subjects in the actual environment and the structure of human skeleton, we first remove part of the outliers by limiting the spatial extent of the body joints. The basis is that in human movement, body joints only move in a certain area and there are specific constraint relationships between the joints. For example, during normal walking, the joints of the left leg will not appear on the right side of the spine. Then we perform K-means clustering on the coordinates of each joints collected by each Kinect instead of all joints. The reason is that when human body occlusion occurs, the joints detected by Kinect will be close to each other or even superposed, which will cause the outliers of a joint to be recognized as normal points of other joints. For the result after clustering, the normal data belongs to the large and dense cluster, while the outlier belongs to the small or sparse cluster, or does not belong to any cluster. Furthermore, a threshold value is set for each cluster according to the Pauta criterion. If the distance from the point in the cluster to the center of the cluster is greater than the threshold value, the point is regarded as an outlier. In general, a point that satisfies any of the following conditions is an outlier: (1) The distance from this point to the nearest cluster center is greater than the threshold value; (2) The cluster to which this point belongs is a small sparse cluster.

After detecting and removing the outliers, we need to unify coordinate systems to facilitate the implementation of data fusion. We adopt the coordinate system transformation algorithm proposed in [19] to convert the 3D coordinates of body joints from different views into Kinect-3's coordinate system.

C. Data Fusion

According to the level of information representation, data fusion can be divided into three categories: data-level fusion, feature-level fusion and decision-level fusion. This paper mainly focuses on data-level fusion, i.e. we directly fuse the data collected by Kinect to obtain more accurate and reliable results. Let P_i denote the coordinate of body joints detected by each Kinect at the same time and W_i is the corresponding weight. Let P be the true value of the measurement result. For the measured values of Kinect sensors, they have statistical independence from each other and each of them is an unbiased estimate of P. Thus, the fusion result, $\hat{P}$, is given by

$$\hat{P} = \sum_{i=1}^{n} W_i P_i \quad (1)$$

$$\sum_{i=1}^{n} W_i = 1 \quad (2)$$

where $i \in \{1,2,\ldots,n\}$, n is the number of Kinect sensors which is 5 in this paper. The total variance σ^2 is

$$\sigma^2 = E\left[(P-\hat{P})^2\right] = E\left[\left(\sum_{i=1}^{n} W_i P - \sum_{i=1}^{n} W_i P_i\right)^2\right] = E\left[\sum_{i=1}^{n} W_i^2 \left(P-P_i\right)^2 + 2\sum_{\substack{i=1,j=1\\ i\neq j}}^{n} \left(P-P_i\right)\left(P-P_j\right)\right] \quad (3)$$

where $E(P)$ is the expectation of the variable P. Because P_i is an unbiased estimate of P, it is easy to figure out

$$E\left(P-P_i\right)\left(P-P_j\right)=0 \quad (4)$$

Thus, the total variance is given by

$$\sigma^2 = E\left[\sum_{i=1}^{n} W_i^2 \left(P-P_i\right)^2\right] = \sum_{i=1}^{n} W_i^2 \sigma_i^2 \quad (5)$$

where σ_i^2 denote the variance of each Kinect sensor. It is easy to see that the total variance is a multivariate quadratic function of the weighting factor and has a minimum value. According to the extremum theory of multivariate functions, the minimum of the total variance and the optimal weighting factor can be obtained as

$$W_i^* = 1 / \left(\sigma_i^2 \sum_{i=1}^{n} \frac{1}{\sigma_i^2}\right) \quad (6)$$

$$\sigma_{\min}^2 = 1 / \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \quad (7)$$

Divide σ_i^2 by $\sigma_{\min}^2$

$$\frac{\sigma_i^2}{\sigma_{\min}^2} = 1 + \sigma_i^2 \sum_{j=1, j\neq i}^{n} \frac{1}{\sigma_j^2} \geq 1 \quad (8)$$

It is obvious that $\sigma_{\min}^2$ is less than the variance of any Kinect sensor. Furthermore, it can be proved that the variance of fusion results obtained by this algorithm is less than that of average weighted fusion algorithms. For each moment in the coordinate sequence of joints, if we take the average of the data collected by multiple Kinects as the fusion result, i.e. the weight is $1/n$. Thus, the variance of the average weighted fusion result, $\bar{\sigma}^2$, is

$$\bar{\sigma}^2 = \sum_{i=1}^{n} \sigma_i^2 / n^2 \quad (9)$$

Compared with the adaptive weighted fusion algorithm, according to Cauchy inequality, we can know

$$\frac{\bar{\sigma}^2}{\sigma_{\min}^2} = \left(\frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2\right)\left(\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\right) \geq \frac{1}{n^2}\left(\sum_{i=1}^{n}\sigma_i \frac{1}{\sigma_i}\right)^2 = 1 \quad (10)$$

which means $\bar{\sigma}^2 \geq \sigma_{\min}^2$. To find the optimal weighting factor, we need to figure out the variance of each Kinect. For the measured values of any two independent Kinect, P_i and P_j, let their measurement errors be e_i and e_j respectively.

$$P_i = P + e_i \quad (11)$$

$$P_j = P + e_j \quad (12)$$

where e_i and e_j are stationary noise with zero mean value. Thus, the variance of Kinect-i is

$$\sigma_i^2 = E\left(e_i^2\right) \quad (13)$$

We define the correlation factors between P_i and P_j, R_{ij}, as $E\left(P_i P_j\right)$, the autocorrelation factor of P_i, R_{ii}, as $E\left(P_i^2\right)$. Since e_i, e_j and P are uncorrelated variables, we can find

$$R_{ij} = E\left(P_i\ P_j\right) = E\left(P^2\right) \quad (14)$$

$$R_{ii} = E\left(P_i^2\right) = E\left(P^2\right) + E\left(e_i^2\right) \quad (15)$$

Thus σ_i^2 can be given by

$$\sigma_i^2 = E\left(e_i^2\right) = R_{ii} - R_{ij} \quad (16)$$

R_{ii} and R_{ij} can be obtained by estimating their time domain values.

$$R_{ii}(k) = \frac{1}{m} \sum_{\substack{t=1 \\ P_i(t)\neq 0}}^{k} P_i(t) P_i(t) \quad (17)$$

where $R_{ii}(k)$ is the value of R_{ii} at the k-th acquisition time, m is the number of non-zero values in the data sequence after k times of collection ($m \leq k$) and $P_i(t)$ is the value of P_i at one acquisition moment, $t \in \{1,2,\ldots,k\}$. If $P_i(t)$ is detected as an outlier, its value will be set to 0, because if outliers are included, e_i and e_j are not stationary noise whose mean value is not zero. This also explains the necessity of outlier detection before data fusion. Similarly,

$$R_{ij}(k) = \frac{1}{l} \sum_{\substack{t=1 \\ P_i(t)\neq 0 \\ P_j(t)\neq 0}}^{k} P_i(t) P_j(t) \quad (18)$$

where $R_{ij}(k)$ is the value of at the k-th acquisition time, l is number of frames in which neither $P_i(t)$ nor $P_j(t)$ is 0 ($l \leq k$). There is a correlation factor between any two Kinect sensors. We take the mean of $R_{ij}(k)$, $\bar{R}_{ij}(k)$, as an estimate of R_{ij}.

$$R_{ij} = \bar{R}_{ij}(k) = \frac{1}{n-1} \sum_{\substack{j=1 \\ j\neq i}}^{n} R_{ij}(k) \quad (19)$$

Thus, we get an estimate of the variance of each Kinect using R_{ii} and R_{ij}. Sequentially, we can figure out the weight of each Kinect and the fusion result. As shown in the equations above, different Kinect sensors can adjust the weight distribution according to the measured values at different moments. The fusion results have linear unbiased minimum variance.

IV. Experiments

The proposed algorithm was implemented in Windows 10 computer system using Visual C# and Python language.

A. Test Cases

To verify the effectiveness of the method, we selected 20 subjects (10 males, 10 females). Five Kinect devices started collecting data synchronously. Each subject walked on a treadmill for one minute and 1,500 frames of each data sequence were taken for the experiment. Each frame contains 3D coordinates of 25 body joints from 5 perspectives. In total, 20×1,500×5×25=3750000 3D coordinates were used for the experiment.

B. Results and Analysis

Based on the 3D coordinates of body joints, we calculated the bone length (the Euclidean distance between adjacent joints) and two common gait parameters. Fig. 3 is the sketch of these gait parameters. The standard deviation is used to evaluate the degree of dispersion of data and the stability of results. To study the effects of each step individually, we compared four processing methods using different settings: (1) RO: remove outliers from initial data collected by each Kinect; (2) AWF: adaptive weighted fuse without removing outliers; (3) RO+AF: remove outliers and average weighted fuse; (4) RO+AWF: remove outliers and adaptive weighted fuse. The length of human bones are theoretically fixed values. If the standard deviation is smaller, the method achieves superior performances. Table 1 lists the results of one subject, which demonstrate that the method proposed in this paper can effectively improve the stability of skeleton data. In fact, our method achieved good experimental results like this for all subjects. Comparing initial data with RO and AWF with RO+AWF, it can be found that outlier detection is effective and can improve the effect of data fusion. Comparing RO+AWF with initial data and RO+AF, it can be found that adaptive weighted fusion can obtain the fusion value with linear unbiased minimum variance and is superior to average weighted fusion.

TABLE I. Standard Deviation (CM) of Bone Length

Bone Length	**Kinect-1**		**Kinect-2**		**Kinect-3**		**Kinect-4**		**Kinect-5**		**AWF**	**RO+ AF**	**RO+ AWF**
	Initial data	**RO**	**Initial data**	**RO**	**Initial data**	**RO**	**Initial data**	**RO**	**Initial data**	**RO**			
SpineShoulder-ShoulderLeft	\[a]	\	2.318	2.149	1.341	1.182	2.286	1.956	2.336	2.073	1.197	1.184	**1.032**
ShoulderLeft-ElbowLeft	\	\	2.142	1.843	1.534	1.488	1.932	1.892	1.296	1.228	1.223	1.537	**1.071**
HipLeft-KneeLeft	3.347	3.194	3.089	2.714	2.508	2.367	2.939	2.741	2.541	2.406	2.611	2.427	**2.243**
KneeLeft-AnkleLeft	3.575	3.347	4.049	3.325	2.683	2.512	3.799	3.357	3.243	2.925	2.596	2.821	**2.434**
AnkleLeft-FootLeft	1.498	1.336	2.063	1.877	\	\	2.111	1.899	1.429	1.338	1.395	1.479	**1.304**
SpineMid-SpineBase	1.016	0.978	0.754	0.707	0.458	**0.442**	0.855	0.789	1.362	1.223	0.622	0.801	0.599
SpineShoulder-ShoulderRight	2.282	1.995	2.879	2.564	1.688	1.593	1.992	1.756	\	\	1.608	1.952	**1.542**
ShoulderRight-ElbowRight	3.038	2.954	1.681	1.459	1.906	1.757	1.314	1.144	\	\	1.261	1.886	**1.182**
HipRight-KneeRight	3.857	3.431	3.142	2.758	2.357	2.208	3.546	3.259	3.266	2.873	2.104	2.586	**1.983**
KneeRight-AnkleRight	3.874	3.469	2.588	2.443	2.329	2.261	3.296	2.842	3.634	3.238	2.558	2.833	**2.077**
AnkleRight-FootRight	1.934	1.669	2.046	1.862	\	\	2.256	1.964	1.787	1.667	1.751	**1.709**	1.713

a. '\' denotes the proportion of outliers in the total data is greater than 1/3.

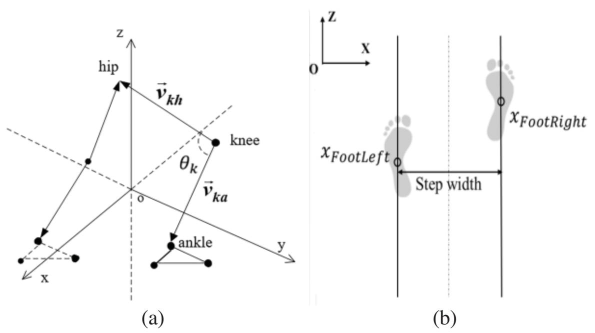

Fig. 3. Sketches of gait parameters. (a) Angle of knee; (b) Step width

We calculated gait parameters: angle of knee and step width to further verify the effectiveness of the algorithm. As shown in Fig. 3, angle of knee, θ_k, can be calculated by

$$\vec{\boldsymbol{v}}_{kh} = \vec{\boldsymbol{v}}_h - \vec{\boldsymbol{v}}_k \tag{20}$$

$$\vec{\boldsymbol{v}}_{ka} = \vec{\boldsymbol{v}}_a - \vec{\boldsymbol{v}}_k \tag{21}$$

$$\theta_k = \cos^{-1}\frac{\vec{v}_{kh}\cdot\vec{v}_{ka}}{|\vec{v}_{kh}|\cdot|\vec{v}_{ka}|} \tag{22}$$

where $\vec{v}_h = (x_h, y_h, z_h)$, $\vec{v}_k = (x_k, y_k, z_k)$, $\vec{v}_a = (x_a, y_a, z_a)$ denote 3D coordinates of hip joint, knee joint and ankle joint respectively. The knee angle is calculated using the left leg as an example in this paper. Since we convert the 3D coordinates of body joints from different views into Kinect-3's coordinate system, step width, L_{sw} , can be calculated by

$$L_{sw} = |X_{FL} - X_{FR}| \tag{23}$$

where X_{FL} and X_{FR} are the x coordinates of FootLeft joint and FootRight joint detected by Kinect respectively. The results of one subject listed in Table 2. In theory, the angle of knee changes periodically during walking.

TABLE II. STANDARD DEVIATION OF GAIT PARAMETERS

Gait Parameters	RO+AF	RO+AWF
Angle of Knee(°)	15.28	**14.4**
Step Width(cm)	2.793	**2.585**

As shown in Fig. 4, compared with the initial data and the results of weighted fusion, the angle of knee obtained by our method has a more obvious periodicity and less noise.

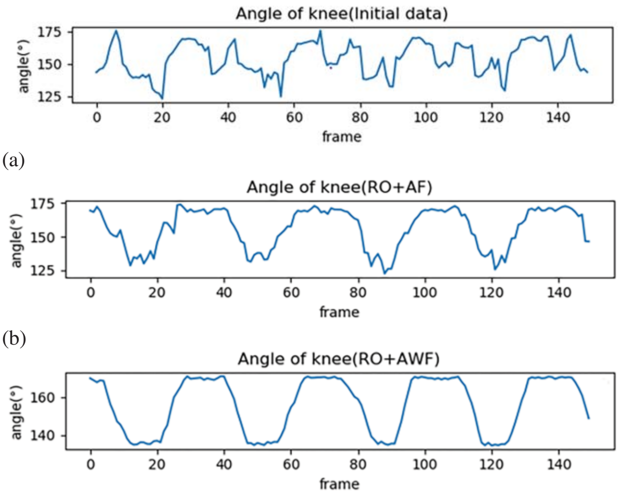

Fig. 4. The change of knee angle during walking. (a) Initial data; (b) RO+AF; (c) RO+AWF.

C. Result Visualization

To show the effect of our method more intuitively, we visualized some experimental results. Fig. 5(a) shows the results of outlier detection on the left knee joint of one subject. Obviously, outliers can be detected and some outliers might gather as small sparse clusters due to human body occlusion, i.e. some joints are wrongly detected as other obscured body joints by Kinect at some moments. Fig. 5(b) and (c) shows the skeleton model drawn from initial data and the results obtained by our proposed data optimization algorithm. It is obvious that invalid data exists in the initial data and our method can obtain more stable and effective skeleton data sequence.

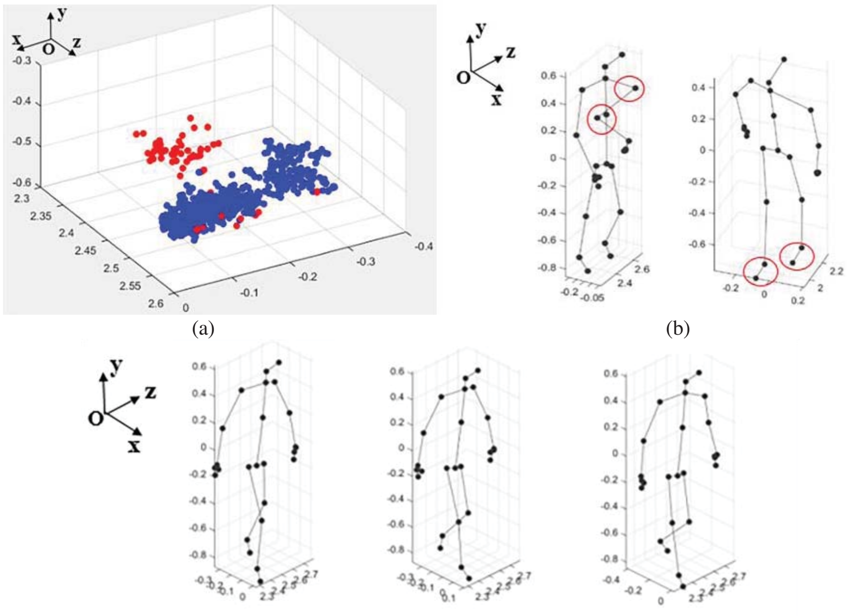

Fig. 5. Result Visualization. (a) The results of outliers detection: the blue points are normal and the red points are outliers; (b) Skelecton model of Initial data. The joints in the red circles are detected outliers; (c) Skelecton model of results.

V. CONCLUSION

We propose an effective multi-view Kinect skeleton data optimization algorithm in this paper, which can process abnormal and wrong 3D coordinates of 25 human joints caused by measuring environment and human body occlusion. Outlier detection can improve the effect of data fusion. The adaptive weighted fusion algorithm can fuse each set of skeleton data collected by different Kinect sensors into a stable and valid data sequence. This method can improve the stability and validity of skeleton data, and make Kinect more easily applied to motion analysis, motion recognition and medical rehabilitation evaluation.

REFERENCES

[1] I. Pachoulakis and K. Kourmoulis, "Building a gait analysis framework for Parkinson's disease patients: Motion capture and skeleton 3D representation," International Conference on Telecommunications and Multimedia, pp. 220-225, 2014.

[2] D. Xu, X. Xiao, X. Wang and J. Wang, "Human action recognition based on Kinect and PSO-SVM by representing 3D skeletons as points in lie group," International Conference on Audio, Language and Image Processing, pp. 568-573, 2016.

[3] S. Sinha, B. Bhowmick, K. Chakravarty, A. Sinha and A. Das, "Accurate upper body rehabilitation system using kinect," Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4605-4609, 2016.

[4] T. C. Hoang, H. T. Dang, V. D. Nguyen , "Kinect-based virtual training system for rehabilitation," International Conference on System Science and Engineering, pp. 53-56, 2017.

[5] Z. Zhang, "Microsoft Kinect Sensor and Its Effect", IEEE MultiMedia, vol. 19, no. 2, pp. 4-10, Feb. 2012.

[6] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, A.Blake, "Real-time human pose recognition in parts from single depth images," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1297-1304, 2011.

[7] K. Khoshelham, "Accuracy analysis of kinect depth data," ISPRS workshop laser scanning, vol. 38, no. 5, pp. 133-138, Aug. 2011.

[8] M. Oliver, F. Montero, J. P. Molina, P. González, A. Fernández-Caballero, "Multi-camera systems for rehabilitation therapies: a study of the precision of Microsoft Kinect sensors." Frontiers of Information Technology & Electronic Engineering, vol.17, no. 4, pp. 348-364, 2016.

[9] D. J. Geerse, B. H. Coolen, and M. Roerdink, "Kinematic validation of a multi-Kinect v2 instrumented 10-meter walkway for quantitative gait assessments," PLoS One, vol. 10, no. 10, pp. 1–15, 2016.

[10] X. Xu, R. W. McGorry, L. S. Chou, J. Lin, and C. Chang, "Accuracy of the Microsoft KinectTM for measuring gait parameters during treadmill walking," Gait Posture, vol. 42, no. 2, pp. 145–151, 2015.

[11] Z. Gan and X. Zhou, "Abnormal Network Traffic Detection Based on Improved LOF Algorithm," International Conference on Intelligent Human-Machine Systems and Cybernetics, pp. 142-145, 2018.

[12] F. Angiulli, S. Basta and C. Pizzuti, "Distance-based detection and prediction of outliers," IEEE Transactions on Knowledge and Data Engineering, vol. 18, no. 2, pp. 145-160, Feb. 2006.

[13] X. Shao, M. Zhang and J. Meng, "Data Stream Clustering and Outlier Detection Algorithm Based on Shared Nearest Neighbor Density," International Conference on Intelligent Transportation, Big Data & Smart City, pp. 279-282, 2018.

[14] M. A. Haque and H. Mineno, "Contextual Outlier Detection in Sensor Data Using Minimum Spanning Tree Based Clustering," International Conference on Computer, Communication, Chemical, Material and Electronic Engineering, pp. 1-4, 2018.

[15] S. Li, P. N. Pathirana and T. Caelli, "Multi-kinect skeleton fusion for physical rehabilitation monitoring," Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 5060-5063, 2014.

[16] H. Du, Y. Zhao, J. Han, X. Zhao, Z. Wang, G. Song, "Data Fusion of Human Skeleton Joint Tracking Using Two Kinect Sensors and Extended Set Membership Filter," Acta Automatica Sinica, vol. 42, no. 12, pp. 1886-1898, Dec. 2016.

[17] Z. Li, R. Chen, and B. Zhang, "Study of adaptive weighted estimate algorithm of congeneric multi-sensor data fusion," Journal of Lanzhou University of Technology, vol. 32, no. 4, pp. 78-82, 2006.

[18] C. B. Pereira, X. Yu, V. Blazek, B. Venema and S. Leonhardt, "Multisensor data fusion for enhanced respiratory rate estimation in thermal videos," Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1381-1384, 2016.

[19] X. Fang, W. Zeng, J. Liu, Y. Yao, "A general total least squares algorithm for three-dimensional coordinate transformations," Acta Geodaetica et Cartographica Sinica, vol. 43, no. 11, pp. 1139-1143, Nov. 2014.