2014 IEEE International Conference on Internet of Things (iThings 2014), Green Computing and Communications (GreenCom 2014), and Cyber-Physical-Social Computing (CPSCom 2014)

Continuous Human Location and Posture Tracking by Multiple Depth Sensors

Jun-Wei Qiu*, Ting-Hui Chiang*, Chi Chung Lo*†, Li-Min Lin*, Lan-Da Van*, Yu-Chee Tseng*, and Yu-Tai Ching*
*Department of Computer Science, National Chiao Tung University, Hsinchu 30010, Taiwan
†Here Co., Ltd., Hsinchu, 30078, Taiwan
Email: {chiucw, tinghui, ccluo, lmlin, ldvan, yctseng, ytc}@cs.nctu.edu.tw;

***Abstract*— This paper proposes a continuous location and skeletal tracking system using multiple RGB-Depth sensors (such as Kinects) deployed along a corridor with overlapping coverages. First, we transform the coordinates of all sensor into a unified coordinate. Second, the system recognizes users (such as patients under rehabilitation) from different views of these sensors and classifies them by their patient IDs. Third, the patient information can be continuously handed over among sensors when they move around the area. Experiment results show that the skeletal association during handover achieves approximately 90.61% in accuracy and 96.89% in precision in 44,817 experiment trials. By our observation, injured people possess asymmetric gait parameters especially on the ratio of the duration of the swing/stance phase. For example, the injured foot generally has a longer swinging duration than the healthy side. The proposed system has potential in patient rehabilitation monitoring applications.**

Keywords: Depth sensor, Gait analysis, Rehabilitation, Sensor network, Skeleton tracking.

I. Introduction

For people suffering from neurological disorder, they typically have limited mobility. Traditionally, rehabilitation and physical therapies are exercised to help their recovery, during which the medical personnels need to monitor such patients to diagnose their posture problems. Nowadays, non-contact sensing technologies, such as RGB-Depth cameras, are promising tools to help track the locations and postures of the patient without the presence of medical personnel. Posture tracking is an interesting but challenging problem introducing various forms of applications such as activity recognition, human-machine interaction, and pose mimicking. To develop an effective solution, both high dimensionality and severe self-occlusions of human behavior need to be conquered. Marker-based optical motion capture (MOCAP) [1, 2] provides a possible solution. However, it requires complex hardware setup in a specific area and target users need to be tagged in advance.

Recently, non-contact RGB-Depth sensors appear as promising sensing tools for posture tracking. For example, Kinect has been successfully applied to the fields of video games, surveillance, and healthcare. The main technology behind the Kinect sensor is depth sensing, which is able to reconstruct 3D skeletons of human bodies using the captured RGB and depth images. However, such RGB-Depth sensors normally have limited coverage per sensor (about 3 to 5 meters by one Kinect). To remove this limitation, this research considers using multiple RGB-Depth sensors for continuous tracking of rehabilitating patients' motions and postures. The achievements of this research are:

- The transformation of multiple sensors' coordinates into one common coordinate system
- The skeleton association during handover among sensors
- Continuous gait analysis with skeletal movements contributed by multiple sensors.

We have performed experiments to verify the feasibility of our system with the Kinect API on Microsoft .net Framework 4.0 with six subjects. Then, we used the recorded coordinates of the skeletal joints to perform gait analysis. Our system can capture and derive gait parameters such as gait length, gait duration and portion of the stance/swing phase within one gait when a user walks inside its covering area. With such capability, we can construct a tracking area and observe people's walking behaviors. Our observation shows that an injured person tends to have a shorter stance phase on the injured side and prolong the supporting duration on the normal side, while a healthy person generally shows symmetric characteristics on both sides.

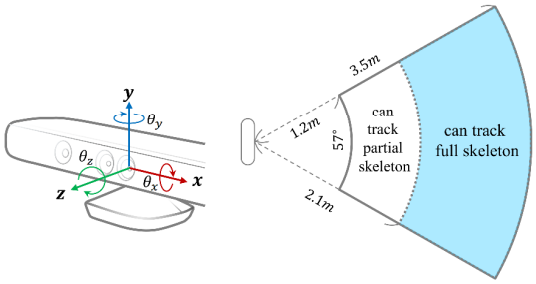
The rest of this paper is organized as follow. Section II presents the related works. Section III introduces our system infrastructure. Section IV shows our experimental results. Section V concludes this paper.

II. Related Works

A. Indoor Activity Tracking

Early development for location tracking includes Active Bat [3] and Cricket [4]. RADAR [5] adopts the Wi-Fi infrastructure for indoor localization and its extensions includes [6, 7]. Inertial sensors are used for activity recognition [8, 9] and pedestrian tracking [10, 11]. Vision-based solutions are discussed in [12, 13]. Later, Kinect was developed as a motion sensing device for Microsoft Xbox, in which an RGB camera, an infrared depth sensor, and a microphone array mounted on a horizontal bar are embedded. It also has the capability of identifying human skeletal structures. Thus, tracking human skeletal structure in its field of view [14] become possible. Recent applications of Kinects are discussed in [15, 16].

978-1-4799-5967-9/14 $31.00 © 2014 IEEE
DOI 10.1109/iThings.2014.31



(a) Axes of a Kinect (b) Tracking area of a Kinect

Fig. 1. Properties of the Kinect sensor.

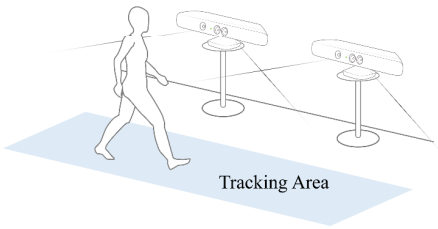


Fig. 2. Illustration of the Scenario.

B. Gait Analysis

Human gait analysis has extensive applications in clinical diagnoses and rehabilitations, such as neurological diseases [17] and fall detection [18]. Marker-based systems are possible solutions, but unfortunately, users have to be tagged when tracking. Force or pressure platforms can observe further information [19], but the sensing field is quite limited and the measurements only contain information from lower limbs. Accelerometer-based sensor arrays provide a cheaper alternative [20], but the test subjects must be tagged and continuous sensor calibration is needed. Advanced vision systems and depth cameras are non-intrusive tools for gait analysis [21, 22]. Kinect is developed based on depth cameras and has been used in biometrics verification [23, 24] and Parkinson's disease analysis [25]. Advanced skeletal tracking applications such as [26] are also proposed to broaden the conventional usage as a gaming console.

III. System Infrastructure

A single Kinect sometimes barely provides sufficient skeletal information for gait analysis, and even misses a human passing by easily. The major reason is that a Kinect has only limited 57° forward-facing horizontal field of view with maximum $3.5m$ in depth. Fig. 1(a) shows the axes of a Kinect sensor and Fig. 1(b) shows the area where the skeleton of a $175cm$-tall subject can be tracked when the Kinect is placed $0.8m$ above the floor. In order to overcome the limitations, our non-contact posture monitoring system contemplates multiple Kinects, as shown in Fig. 2, over a corridor to track and record the locations of 20 skeletal joints continuously when the subjects enter the tracking area.

In our scenario, the system should be able to monitor the subjects' gaits. It can handle handover events continuously when the subjects move from the field of view of one Kinect to another, and observe gait parameters in the tracking duration. Skeletal tracking is one of the popular way to achieve the goal of gait analysis, as most of the abnormalities can be detected by the captured joint motions of the lower limbs.

Theoretically, a single Kinect sensor can track a full skeleton without clipping in a $2.8m$ straight line with its field of view, but it requires the user's face, upper body and pre-processing time for the Kinect sensor to successfully identify and track a skeleton. Therefore, the actual skeletal tracking region is much narrower than the field of view described in specifications while the tracked subjects pass by the field of view instead of staying in it. In order to track skeletons robustly, visualize their locations and identify postures correctly, we develop an indoor tracking service by using multiple Kinects with seamless handover capabilities so that the coverage limitation can be alleviated. To develop the multiple-Kinect infrastructure, we should consider the following issues:

A. Coordinate Transformation among Multiple Kinects

When a user enters the tracking area, the system starts tracking the skeleton automatically and continuously until the user exits the tracking area. The recorded skeletal frames are sent to the data collecting server via LAN connections, the synchronization of the recorded frames is done by merging the Kinect time stamp of the recorded frame on the server side. To combine the skeletons detected by multiple Kinects in overlapping fields of view, a common coordinate system of multiple Kinects should be established to ensure the measurements of one Kinect can be successfully identified and handed over to another. Each coordinate of a Kinect can be transformed to the common coordinate using the rotation matrix derived from the orientation and the displacement of the device.

Let $(\theta_x^i, \theta_y^i, \theta_z^i)$ be the orientation of a Kinect device K^i around the three axes of the common coordinate, and $(\Delta x^i, \Delta y^i, \Delta z^i)$ be the displacements of the device from $(0, 0)$. Let $S^i = \{u_n^i | n = 1, ..., 20\}$ be the skeleton detected by Kinect K^i, where $u_1^i, u_2^i, ..., u_{20}^i$ be the skeletal joints. Each joint coordinate u_n^i of the Kinect K^i can be converted into common coordinate using the following transformation, denoted as $T_{K^i}(\cdot)$:

$$T_{K^i}(u_n^i) = R_z(\theta_z^i)R_y(\theta_y^i)R_x(\theta_x^i) \cdot u_n^i + \begin{bmatrix} \Delta x^i \\ \Delta y^i \\ \Delta z^i \end{bmatrix} \quad (1)$$

where the three rotation matrices are:

$$R_z(\theta_z^i) = \begin{bmatrix} \cos\theta_z^i & -\sin\theta_z^i & 0 \\ \sin\theta_z^i & \cos\theta_z^i & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$$R_y(\theta_y^i) = \begin{bmatrix} \cos\theta_y^i & 0 & \sin\theta_y^i \\ 0 & 1 & 0 \\ -\sin\theta_y^i & 0 & \cos\theta_y^i \end{bmatrix} \quad (3)$$

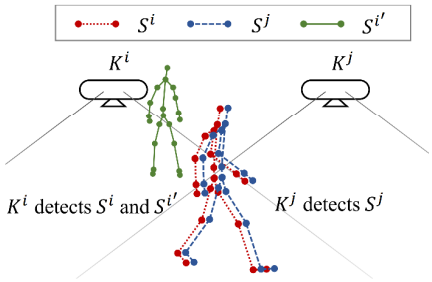
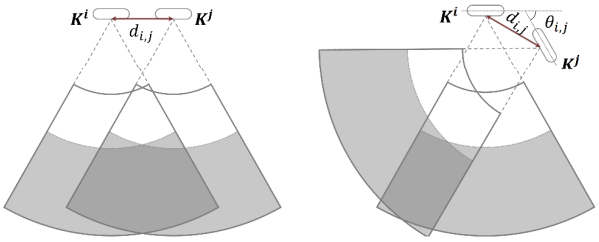


Fig. 3. An example of association, S^i and S^j are associated if $\epsilon_{i,j} < d_{th}$.



(a) Parallel deployment

(b) Two Kinects with θ differences in viewing angle

Fig. 4. Deployment of Kinect sensors.

$$R_x(\theta_x^i) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x^i & -\sin\theta_x^i \\ 0 & \sin\theta_x^i & \cos\theta_x^i \end{bmatrix} \quad (4)$$

The Kinect sensor can provide a three-dimensional virtual skeleton consisting of 20 coordinates of skeletal joints at the sampling rate of 30 frames per second.

B. Skeleton Association for Handover

Since there might be several people moving around in the tracking region simultaneously, and the tracking region is covered by multiple Kinects with overlapping fields of view. The system infrastructure needs to identify whether the skeletons detected by different Kinects belong to the same person or not. The observed coordinates of the skeleton joints are transformed and compared to make decisions. However, some joints might be obstructed during the tracking; therefore, the skeleton might not consist of a complete 20 fully tracked joints, and some joints are only inferred or become not available at all.

Each Kinect can detect two skeletons simultaneously, we denote two skeletons detected by Kinect K^i as S^i and $S^{i'}$. When a skeleton frame $S^i = \{u_n^i | n = 1, 2, ..., 20\}$ is detected by Kinect K^i, the skeletal joints are partitioned into three subsets J_{trk}^i, J_{inf}^i and J_{non}^i according to their visibility in the field of view of K^i. The tracked set J_{trk}^i includes the coordinates of the tracked joints, the inferred set J_{inf}^i contains the coordinates of the inferred joints, and those which belong to the non-tracked set J_{non}^i do not have sufficient position information and will be discarded in our design. Let K^i, K^j be the two Kinects and both detect a single skeleton in their field of view, say $S^i = \{u_1^i, u_2^i, ..., u_{20}^i\}, S^j = \{u_1^j, u_2^j, ..., u_{20}^j\}$, where u^i and u^j are three-dimensional vectors denoting the positions of the tracked skeleton joints.

Let $C_{i,j} = \{n | u_n^i \in J_{trk}^i, u_n^j \in J_{trk}^j\}$ be the subset of common tracked joints detected by both K^i and K^j, and $\epsilon_{i,j}$ be the averaged joint position error, which is defined as:

$$\epsilon_{i,j} = \frac{1}{|C_{i,j}|} \cdot \sum_{n \in C_{i,j}} ||T_{K^i}(u_n^i) - T_{K^j}(u_n^j)||, \quad (5)$$

where $T_{K^i}(u_n^i)$ denotes the transformed location of joint coordinate u_i detected by Kinect K^i in the common coordinate system. The system associates two skeletons S^i and S^j as the same person if $\epsilon_{i,j} \leq d_{th}$.

Fig. 3 illustrates an example of the association. Kinect K^i detects two skeletons S^i and $S^{i'}$ and Kinect K^j detects S^j. The system will calculate associations for $\{S^i, S^j\}$ and $\{S^{i'}, S^j\}$. If $\epsilon^{i,j} < d_{th}$, the system will associate skeleton S^i and S^j as the same person, instead of $S^{i'}$ and S^j. If both $\epsilon^{i,j}$ and $\epsilon^{i',j}$ are smaller than d_{th}, then the system will associate the one with S^j that creates the smallest difference.

To determine the proper distance threshold d_{th} for association, we have performed the corresponding experiments using two Kinects at various distances $d_{i,j}$ and different viewing angles $\theta_{i,j}$. The deployment of two Kinects system is illustrated in Fig. 4. In our experiments, a single subject is told to walk in front of two Kinects ranging each other from 1, 1.5 and 2 meters, in parallel deployment ($\theta_{i,j} = 0°$) illustrated in Fig. 4(a), and at different viewing angles of $\theta_{i,j} = 30°$ and $60°$ as shown in Fig. 4(b). Thus, total nine different configurations can be obtained.

The testing field is a $10m \times 10m$ indoor area with sufficient illumination approximately at 400lx. The data set contains ten recording trials of ten different subjects for each of nine configurations. Each recording trial elapses for about two minutes and approximately 1,500 pairs of skeletal frames are tracked simultaneously by both two Kinects in one configuration.

The results are shown in Fig. 5, where we group the frame pairs by the number of the tracked common joints $|C_{i,j}|$ observed by K^i and K^j. The column chart shows the average joint distance error $\epsilon_{i,j}$, and the area chart represents the number of frame pairs containing the corresponding number of tracked common joints. We set the association distance threshold $d_{th} = 25cm$, as the average distance errors of the tracked common joints never exceeds $25cm$ for a single person moving in the coverage region in all configurations.

C. Gait Analysis with Skeletal Movements

By using the designed association rule and the handover function, we create a rectangular tracking area with the Kinect array to track the locations of skeletal joints continuously. To get the key parameters for gait analyses, two types of information are required:

- The swinging traces of both foot joints.
- The time of foot strike events on both feet.

The swinging trace of a foot in the common coordinate system can be obtained by merging the coordinates of the

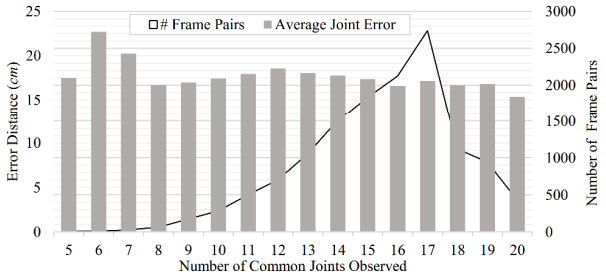


Fig. 5. Average error distance $\epsilon_{i,j}$ and the number of frame pairs by different numbers of tracked common joints $|C_{i,j}|$.

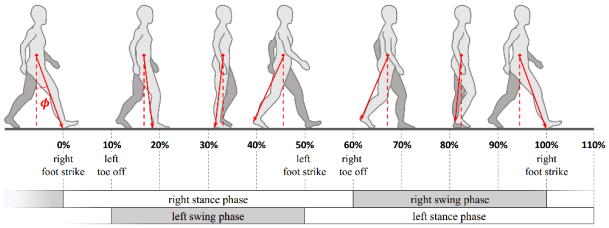


Fig. 6. Illustration of a normal human gait cycle.

foot joints from the corresponding side of the Kinect system. Fig. 6 illustrates a normal human gait cycle. The stance phase of a single foot starts with the corresponding heel striking the ground and ends with the toes leaving it. We define ϕ to be the angle between the segment between ankle to body center, and the segment between body center and its projection to the ground, which is assumed to be perfectly horizontal. In this research, the foot strike events are detected by finding the local maximums of angle ϕ.

To avoid obstructions, we deploy two pairs of Kinect systems to track both left and right sides of the subject's skeleton separately. Then, we use the foot strike events to extract the movements of both feet in the entire swinging trace. When a maximum is identified as a foot strike event, the succeeding maximum will be the event that the ankle is swung behind the body, so the next actual foot strike event will be the next other maximum of ϕ. With the recorded sequence of joint coordinates, ϕ can be derived from the motion of skeletal joints recorded by the Kinect pairs on both sides. Fig. 7 shows the change of ϕ of the right foot ankle on a normal subject in the duration of two gait cycles.

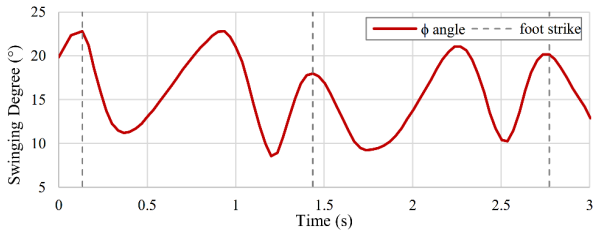


Fig. 7. Variation of ϕ during period of two right gaits.

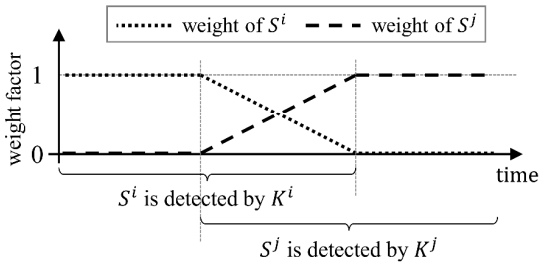


Fig. 8. Variation of ϕ during period of two right gaits.

To obtain a continuous stream of joint coordinates on each side, the Kinect pair on each side weights their coordinate during the handover period. When the first Kinect K^i detects S^i, the weight of the coordinate provided by K^i is 1. After S^j is detected by K^j and the system associates S^i and S^j as the same person, the system will mark the start of the handover period. When S^i disappears, the system will mark the end of the handover period. The weight of the coordinate of S^i is linearly decreased to 0 during the period, while the weight of the coordinate of S^j is linearly increased to 1. The weight function is illustrated in Fig. 8.

Let $\Gamma_\ell[n], n = 1, ..., N$ and $\Gamma_r[m], m = 1, ..., M$ be the ankle movement curves within one gait cycle of the left and right sides, respectively. The step vectors v_ℓ and v_r of the left and right feet can be calculated as $v_\ell = \Gamma_\ell[N] - \Gamma_\ell[1]$ and $v_r = \Gamma_r[M] - \Gamma_r[1]$. From the two curves and two vectors, we can derive the following key terms for gait analysis:

- Step length: the distance between corresponding successive points of heel contact of the opposite feet, which can be calculated from $||v_\ell||$ and $||v_r||$.
- Swinging area: the areas (denoted as A_ℓ and A_r) which are bounded by Γ_ℓ, v_ℓ and Γ_r, v_r.
- Gait duration: the time duration of the left step and the right step (denoted as t_ℓ and t_r), which can be determined using the successive foot strike event on the same foot.
- Time duration of the swing/stance phase: the time duration of the swing/stance phase can be determined using the foot strike event and the toe off event. The portion of the swing phase over the entire gait duration of both feet are denoted as $r_{\ell,sw}$ and $r_{r,sw}$.

In our experiments, different subjects are told to walk in straight line into the tracking area covered by four Kinects one by one. These terms will be calculated and observed using the recorded frames with foot strike and toe off event detection mechanisms.

IV. Experiment Results

Two sets of experiments are performed to verify the association rules described in Section III-B and the gait analysis method proposed in Section III-C.

A. Accuracy of the Skeletal Association Rule

In a common rehabilitation environment, there's usually a medical assistant moving beside the patient to prevent

TABLE I
PERFORMANCE OF THE ASSOCIATION RULE.

	same subject	different subjects
associated	19,703 (43.96%)	1,988 (4.44%)
not associated	2,221 (4.96%)	20,905 (46.65%)

accidents or provide medical care or suggestions. This implies that our system may operate under the circumstances that there are multiple people in the monitoring region. We simulate this environment by ordering two subjects to move simultaneously in the coverage area of the Kinect tracking system. The two subjects are told to wear clothes with contrasting colors. In our case, one is red and another is cyan.

To identify whether the associated skeletons belong to the same subject, the system will examine the hue value of the majority color on the skeletal joints covered by clothes after a pair of skeletons are associated as a single person. If the hue of the majority joint color resides within $90°$ and $270°$, the system classifies the color as cyan, otherwise, as red.

We perform a set of experiments to observe the ratio of correct associations. The Kinects are deployed in parallel, ranging each other from $1.5m$. Two subjects wearing clothes in contrasting colors are told to move around in the coverage area for about 1.5 hours. The system applies the association rule to the tracked skeletons as described in Section III-B, and we use the color as the ground truth of correctness. If the skeletons of the same subject appears in both Kinect coverage areas but no association is occurred for $500ms$, we consider the system fails to associate the skeletons.

We repeat the experiment trial ten times, where total $44,817$ usable trials which at least one skeleton is detected by the Kinects are extracted. The results are shown in Table I, we classify our association results into four cases:

- *True positive (TP):* the associated skeletons belong to the same person in reality.
- *True negative (TN):* skeletons belong to different people are not associated.
- *False positive (FP):* the associated skeletons belong to different people in reality.
- *False negative (FN):* skeletons belong to the same person are not associated.

The terms of accuracy and precision are defined as:

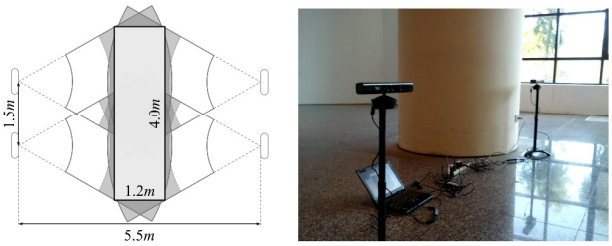
$$\text{accuracy} = \frac{|TP \cup TN|}{|TP \cup TN \cup FP \cup FN|} \tag{6}$$

$$\text{precision} = \frac{|TP|}{|TP \cup FP|} \tag{7}$$

Following the association rule, 90.61% accuracy and 96.89% precision can be achieved via the developed multiple Kinects tracking system.

B. Observation on Gait Parameters

In a $10m \times 10m$ testing field, we deploy two pairs of Kinect systems facing the opposite directions to create a rectangular coverage area. The two Kinect systems in each pair are deployed in parallel, ranging each other from $1.5m$, and the distance between the opposite system is $5.5m$. This configuration creates a tracking area approximately $4m$ long and $1.2m$ wide, which is illustrated in Fig. 9(a).



(a) Deployment of Kinects (b) One of the Kinect pairs

Fig. 9. The configuration of two Kinect system pairs for gait analysis.

TABLE II
GAIT ANALYSIS PARAMETERS OF SIX SUBJECTS.

	$N1$	$N2$	$N3$	$RK4$	$LK5$	$RA6$
$\|\|v_\ell\|\|$ (m)	1.15	0.80	1.17	1.06	0.79	1.01
$\|\|v_r\|\|$ (m)	1.10	0.79	1.28	0.72	0.24	0.31
A_ℓ (m^2)	0.147	0.086	0.305	0.579	0.017	1.079
A_r (m^2)	0.153	0.095	0.267	0.170	0.064	0.176
t_ℓ (s)	1.86	1.80	1.73	2.26	2.47	1.94
t_r (s)	1.81	1.67	1.67	2.87	1.90	2.73
$r_{\ell,sw}$ $(\%)$	41.40	44.00	43.99	20.72	54.81	43.18
$r_{r,sw}$ $(\%)$	39.71	45.53	45.36	75.51	35.09	52.34
height (m)	1.80	1.65	1.74	1.73	1.70	1.75

We recruit six subjects to take an $8m$ round trip within the tracking area. In each recording session, a subject is told to walk into the tracking area covered by four Kinects.

Among the six subjects, $N1, N2$ and $N3$ are healthy; $RK4, LK5$ and $RA6$ have injures on their right knee, left knee and right ankle, respectively. The results in Table II show that the injured subjects possess obvious asymmetric gait parameters in stride length ($||v_\ell||, ||v_r||$), swinging area (A_ℓ, A_r), and the ratio of swing and stance phases ($r_{\ell,sw}, r_{r,sw}$) on the left and right feet comparing to the healthy subjects. According to our observation during the experiments, subjects $RK4, LK5$ and $RA6$ tend to put less pressure on their injured foot by decreasing the time duration to support the body weight. Such characteristic is reflected directly on the swing/stance phase portion over the entire gait. The normal subjects have approximately the same swing/stance phase portion on both feet. However, among the injured subjects, the injured feet can be identified easily as the injured foot has significantly larger swing phase portion over 50% comparing to the healthy side in the whole gait cycle.

V. CONCLUSIONS

In this research, we propose a non-contact tracking system using multiple Kinects to continuously track human skeletons, which diminishes the problem of the limited field of view of a single Kinect. We devise a skeleton association rule and implement handover function to communicate the skeletons detected by different Kinects to determine whether they belong to the

same person. A set of experiment with different viewing angles and inter-Kinect range are evaluated to help us design the parameters of the association rule. We perform experiments with six subjects to check the successful rate of associations. The configuration achieves about 90.61% in accuracy and 96.89% in precision in $44,817$ usable experiment trials. Finally, we use two pairs of Kinect systems to perform gait analysis with the continuous joint coordinates provided on both sides of the tracking area, the injured subjects who possess asymmetric gait parameters, especially with the swing/stance phase portion, can be identified easily with our system.

ACKNOWLEDGMENT

Y.-C. Tseng's research is co-sponsored by MoE ATU Plan, NSC 101-2221-E-009-024-MY3, NSC 102-2218-E-009-002, NSC102-2911-I-002-001, NTU103R7501, IVF-NSC joint research grant 21280013 (102-2923-E-009-001-MY2), Academia Sinica AS-102-TP-A06, ITRI, hTC, Delta, and D-Link.

REFERENCES

[1] "Optitrack motion capture system." [Online]. Available: http://www.naturalpoint.com/optitrack/

[2] "Vicon motion capture solutions." [Online]. Available: http://www.vicon.com/

[3] A. Harter, A. Hopper, P. Steggles, A. Ward, and P. Webster, "The anatomy of a context-aware application," *Wireless Networks*, vol. 8, no. 2/3, pp. 187–197, 2002.

[4] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, "The cricket location-support system," in *Proc. ACM/IEEE MobiCom*, 2000, pp. 32–43.

[5] P. Bahl and V. N. Padmanabhan, "RADAR: An In-building RF-based User Location and Tracking System," in *Proc. IEEE INFOCOM*, vol. 2, 2000, pp. 775–784.

[6] T. Roos, P. Myllymäki, H. Tirri, P. Misikangas, and J. Sievänen, "A probabilistic approach to wlan user location estimation," *International Journal of Wireless Information Networks*, vol. 9, no. 3, pp. 155–164, 2002.

[7] A. Savvides, C.-C. Han, and M. B. Strivastava, "Dynamic Fine-Grained Localization in Ad-Hoc Networks of Sensors," in *Proc. ACM/IEEE MobiCom*, 2001, pp. 166–179.

[8] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive Computing*, 2004, pp. 1–17.

[9] R. Zhu and Z. Zhou, "A real-time articulated human motion tracking using tri-axis inertial/magnetic sensors package," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 12, no. 2, pp. 295–302, 2004.

[10] E. Foxlin, "Pedestrian tracking with shoe-mounted inertial sensors," *IEEE Computer Graphics and Applications*, vol. 25, no. 6, pp. 38–46, 2005.

[11] C.-W. Tan and S. Park, "Design of accelerometer-based inertial navigation systems," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 6, pp. 2520–2530, 2005.

[12] Q. Cai and J. K. Aggarwal, "Tracking human motion in structured environments using a distributed-camera system," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 11, pp. 1241–1247, 1999.

[13] S. L. Dockstader and A. M. Tekalp, "Multiple camera tracking of interacting and occluded human motion," *Proc. IEEE*, vol. 89, no. 10, pp. 1441–1455, 2001.

[14] M. R. U. Saputra, G. D. Putra, P. I. Santosa *et al.*, "Indoor human tracking application using multiple depth-cameras," in *Proc. IEEE Int. Conf. on Advanced Comput. Sci. and Inform. Syst.*, 2012, pp. 307–312.

[15] T. Dutta, "Evaluation of the kinect sensor for 3-d kinematic measurement in the workplace," *Applied ergonomics*, vol. 43, no. 4, pp. 645–649, 2012.

[16] 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.

[17] N. L. Keijsers, M. W. Horstink, and S. C. Gielen, "Ambulatory motor assessment in parkinson's disease," *Movement Disorders*, vol. 21, no. 1, pp. 34–44, 2006.

[18] J. M. Hausdorff, D. A. Rios, and H. K. Edelberg, "Gait variability and fall risk in community-living older adults: a 1-year prospective study," *Archives of physical medicine and rehabilitation*, vol. 82, no. 8, pp. 1050–1056, 2001.

[19] R. A. Clark, A. L. Bryant, Y. Pua, P. McCrory, K. Bennell, and M. Hunt, "Validity and reliability of the nintendo wii balance board for assessment of standing balance," *Gait & posture*, vol. 31, no. 3, pp. 307–310, 2010.

[20] C.-C. Yang and Y.-L. Hsu, "A review of accelerometry-based wearable motion detectors for physical activity monitoring," *Sensors*, vol. 10, no. 8, pp. 7772–7788, 2010.

[21] M. Nixon *et al.*, "Automated markerless analysis of human gait motion for recognition and classification," *ETRI J.*, vol. 33, no. 3, pp. 259–266, 2011.

[22] L. Igual, À. Lapedriza, and R. Borràs, "Robust gait-based gender classification using depth cameras," *EURASIP J. on Image and Video Process.*, no. 1, pp. 1–11, 2013.

[23] Y. Manabe, R. Saito, and K. Sugawara, "Biometric gait verification by horizontal swings in frontal manner towards human-aware environment," in *Proc. IEEE Int. Conf. on Cognitive Informatics & Cognitive Computing*, 2012, pp. 219–225.

[24] Y. Manabe and K. Sugawara, "Soft biometric verification by integrating static and dynamic features based on fuzzy inference," in *Proc. IEEE Int. Conf. on Cognitive Informatics & Cognitive Computing*, 2013, pp. 242–247.

[25] S. Pohoaţă, O. Geman, and A. Graur, "Dual tasking: gait and tremor in parkinsons disease–acquisition, processing and clustering," in *Proc. National Symposium of Theoretical Elect. Eng.*, 2012.

[26] D. S. Alexiadis, P. Kelly, P. Daras, N. E. O'Connor, T. Boubekeur, and M. B. Moussa, "Evaluating a dancer's performance using kinect-based skeleton tracking," in *Proc. ACM Int. Conf. on Multimedia*, 2011, pp. 659–662.