Tai-Chi Chuan Motion Tracking by Integrating Inertial and Monocular Visual Sensors

Cheng-Ta Shen
CSIE, NCNU
chengtashen@gmail.com

Nai-He Hsu
CSIE, NCNU
ksbcboy@gmail.com

Chih-Yi Chiu
CSIE, NCYU
cychiu@iis.sinica.edu.tw

Sheng-Wen Shih
CSIE, NCNU
swshih@ncnu.edu.tw

***Abstract*—In this paper, we study the 3-D human posture tracking problem for computer-aided Tai-Chi Chuan motion learning. A model-based human motion tracking method is proposed which uses a monocular visual sensor and multiple accelerometers attached to different body parts of a subject. In order to track human postures robustly, a generic *complete and parametrically continuous* (CPC) kinematic model is constructed. Twelve CPC parameters are adapted according to the lengths of the upper limbs and the joint positions of the shoulders and the thighs so that the generic CPC model can be used to track different subjects. Equations for predicting the acceleration measurements are derived using the kinematic model. A method for assessing the difference between the measured and the predicted silhouette images is provided. The particle filter is adopt to track human posture from torso toward the ends of the limbs. Real experiments have been conducted to test the proposed method with various human motions including five simple actions and a fundamental Tai-Chi Chuan actions (i.e., "ward off", "roll back", "press", and "push"). The results show that the proposed method is very promising.**

***Index Terms*—Human Motion Tracking, Accelerometer, Particle Filter.**

I. Introduction

Tracking human motion has been an important topic in computer vision because many practical applications can benefit from the human motion information. In particular, tracking human motion is an integral part to developing a computer-aided learning system for sports. In this work, we will study the motion tracking problem of computer-aided learning for Tai-Chi Chuan. Tai-Chi is a popular exercise which is good for health and is suitable for people of all ages including postoperative and/or asthma patients. Although learning Tai-Chi looks easy, every Tai-Chi action is composed of complex movements and it might take decades to get familiar with the whole action. Beginners often need an experienced coach to assist learning. Chua *et al.* developed an immersion Tai-Chi learning environment [1] in their MasterMotion project using a motion-capture (mocap) device. The mocap data of the student and the Tai-Chi teacher are compared and the student error is displayed in their virtual reality (VR) system. Through this visual feedback, the student can correct their Tai-Chi motion accordingly. However, the cost of the mocap device and the VR environment is too high to popularize their learning system, not to mention that users have to wear a head-mounted display and a bodysuit with dozens of optical tracking balls. Therefore, a computer-aided Tai-Chi learning system is demanding a less-restrictive, low cost and efficient motion tracking method. The purpose of human motion tracking is to use different kinds of sensors, such as inertial sensors, cameras, magnetics and ultrasound systems [2–4], to estimate the orientations and positions of human body parts. In recent years, the vision-based motion capture has become a popular research topic. A visual sensor can provide shape, color, and location information for posture tracking. However, cluttered background, variation of the subject body shape, different clothing, and self-occlusion will all make visual tracking of the human posture a very challenging problem. Additionally, the high degrees of freedom (DOFs) of a human body constitute a high dimensional state space which makes searching the optimal state a very time-consuming problem. In a computer-aided learning system, the cluttered background problem can be alleviated by training a background model. Moreover, the clothing can be designed to improve the tracking accuracy. While the self-occlusion problem can be alleviated by using multiple cameras [5–10], the cost of setting up a multiple camera network is proportional to the number of cameras used. Therefore, to reduce the number of cameras in human motion tracking, cost-effective accelerometers have been used to provide supplemental information to track the human posture [11–16]. Accelerometers not only can measure the dynamics of a moving body but also can provide partial body orientation information from the acceleration of gravity. As to the high DOFs problem, a common solution [17] is to divide the posture tracking problem into many sub-problems of tracking different body parts. Since the DOFs of each body part is much less than that of the whole body, the computation complexity of the posture tracking problem can be dramatically reduced. Furthermore, measurements from the inertial sensors are useful to reduce the search space of each body part. Estimation of the human posture is usually accomplished using the particle filtering technique [18] because of its effectiveness in handling nonlinearity and multiple hypotheses.

In this work, a human posture tracking method is proposed which uses a single camera and multiple accelerometers attached to different body parts of a subject. Notably, it is difficult to track human motions using a monocular camera because of occlusion and the lack of depth information. To simplify the 3-D posture tracking problem in a monocular video sequence, a model-based tracking approach is adopted. A generic *complete and parametrically continuous* (CPC) kinematic model [19] of a human body is created. To fit the generic CPC model to different subjects, 12 CPC parameters are adapted according to the lengthes of the upper limbs and the joint positions of the shoulders and the thighs. The 12 parameters, 20 joint angles, and the six motion parameters of the torso have to be estimated in the tracking process. The particle filter technique is used to track the body parts in a top-down sequence from the torso toward the ends of the limbs.

The remainder of this paper is organized as follows. Section II describes the related work. Section III introduces the sensors adopted in this work, the CPC kinematic model of the human body, and the proposed top-down method for 3D human posture tracking. Section IV shows the experimental results of the proposed method. Conclusions and future work are described in Section V.

II. Related Work

Thanks to the advances of the micro electro-mechanical system (MEMS) techniques, miniature inertial sensors, such as rate-gyros and accelerometers, are gaining popular in different applications. They also have been adopted in studying human motions. A rate-gyro/accelerometer is usually attached to an object to estimate its orientation/inclination from the angular velocity/acceleration measurements. Existing methods of motion analysis using inertial sensors can be classified into the following three categories according to their applications.

1) Human-computer interaction: The accelerometers have been adopted in many devices, such as game consoles and cellular phones, to assist human-computer interaction. The inertial-sensor-based interface can be made very sophisticated. For example, Liu *et al.* [11] implemented a wearable sensor system consisting of two-axis accelerometers and rate-gyros to control a humanoid robot.
2) Motion analysis without posture reconstruction: Junker *et al.* [14] developed a gesture spotting system with body-worn inertial sensors to detect user activities and exploited the recognition capabilities of HMM models to classify the gestures. Ward *et al.* described an activity recognition method using two sensor modules attached on the wrist and the upper arm, respectively [16]. Their sensor module consists of a microphone and an accelerometer. Each activity is characterized by a hand motion and an accompanying sound. For instance, hammering is portrayed by the rise and fall of the arm, accompanied on impact by a loud bang. Also, the usefulness of inertial sensors have been proved in computer-aided learning system for sports and/or musical instruments. Chi *et al.* [20] used piezoelectric force sensors to support the judges in scoring valid attacks in taekwondo competition sparring. Takahata *et al.* [21] also applied accelerometers to karate training. The accelerometer data assists a coach to assess the difference between a beginner and an expert by measuring the rhythm of motion. Fong *et al.* [13] develop an acceleration-based

wireless upper limb motion sensing system for recording the hand motion of a running subject in a sports science application. Kunze *et al.* use body-worn gyroscopes and accelerometers to analyze the smoothness of Tai-Chi Chuan movements [22]. They proposed to assess the energy consumption of a Tai-Chi practicer using the energy of the measured inertial signals. Since a Tai-Chi expert usually consumes less energy than a beginner when playing the same Tai-Chi action, their system was trained to classify the grade of a Tai-Chi practicer using the estimated energy consumption. Spelmezan and Borchers [23] proposed a wireless computer-aided system for real-time snowboard training. The proposed system consists of two Bluetooth Shake SK6 inertial sensor packs, two bend sensors, and force-sensitive resistors. The sensors are used to detect common mistakes of beginner snowboarders, such as insufficient knee bending, incorrect weight distribution, and incorrect rotation of the upper body, and are fed back to the snowboarders immediately. Kwon and Gross [24] proposed a martial art training system using a visual sensor and a wrist-mounted wireless accelerometer. Their system monitors and evaluates the motion of a trainee's single arm and then generates a visual feedback for gesture training. Schoonderwaldt *et al.* [25] develop a cost-effective method combining a two-axis accelerometer and a video camera for estimating the bow velocity in violin playing.

3) Motion analysis with posture reconstruction: Luinge and Veltink described a preliminary study of using the Kalman filter technique to estimate the inclination of a human torso by considering the dynamics of the torso [12]. Mayagoitia *et al.* [26] developed an optical motion analysis system combining accelerometers and rate-gyros to estimate the joint angles, velocities and accelerations of lower limbs in the sagittal plane. The accuracy of the proposed method is verified using a mocap system. Zhu and Zhou [15] propose a real time tracking method using sensor modules to track human motion. Each sensor module consists of a tri-axis accelerometer, a tri-axis rate-gyro and a tri-axis magnetic sensor. Fifteen sensor modules are attached to different body parts (12 modules for the four limbs and three modules for the torso). Since the magnetic sensors can provide the position and orientation measurements, the body posture can be readily estimated using a linear Kalman filter.

Among the aforementioned methods, the one proposed by Zhu and Zhuo [15] is the most similar to ours in terms of using multiple body sensors to track the whole body posture. However, metals in the environment may distort the measurements of a magnetic sensor. Also, the orientation angle obtained by integrating a rate-gyro signal tend to drift over time. Using magnetic sensors and rate-gyros will not only increase the system cost but also introduce more uncontrollable distortions and measurement biases. Therefore, in our system, only tri-axis accelerometers are included as the body sensors. Additionally, a monocular visual sensor is used to provide supplemental information for estimating the whole body posture. The measurements from both the visual sensor and the body sensors are provided to particle filters for estimating the human motion parameters. Fig. 1 shows an overview of our method, which will be described in depth in the following sections.

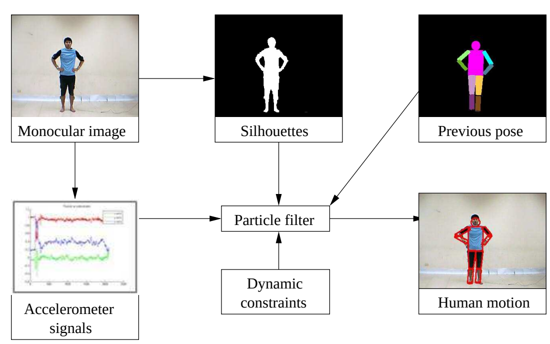


Fig. 1. Overview of the proposed method.

III. Human Motion Tracking

A. Body Sensors Schematic

In this work, nine tri-axis MEMS accelerometers are attached to nine body parts, which are denoted by torso, left upper arm (LUA), right upper arm

(RUA), left forearm (LFA), right forearm (RFA), left thigh (LT), right thigh (RT), left shank (LS), and right shank (RS), respectively, as shown in Fig. 2. At present, the body sensor network are connected by cables to a microprocessor with a star-graph topology. The acceleration signals are digitized with a 12-bit resolution. The microprocessor transmits 30 samples per second to the host computer, where each sample is a 27-D vector. The body sensor signals are interpolated and then are re-sampled such that the body sensor signals are synchronized with the video.

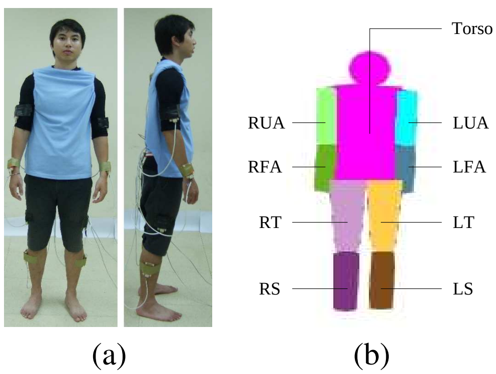


Fig. 2. (a) Body sensor placement. (b) A generic human body model.

B. The Kinematic Model of the Human Body

The generic human body model used in this work consists of a head sphere, eight limb cylinders, and a torso cuboid, as shown in Fig. 2(b). Since human body is an articulated object, each body part has a pivot listed in Table I.

TABLE I
PIVOTS OF THE BODY PARTS

Body Part	Pivot Position
Torso	5th lumbar vertebrae [27]
LUA	Left Shoulder
RUA	Right Shoulder
LFA	Left Elbow
RFA	Right Elbow
LT	Left Hip
RT	Right Thigh
LS	Left Knee
RS	Right Knee

The relationships among the joint angles and the pose of each body part can be described by a kinematic model. In order to formulate the coordinate systems of the body sensors, we need a complete (i.e. a 6-DOFs) kinematic model. Therefore, the complete and parametrically continuous (CPC) kinematic model [19, 28] is adopted. Let ${}^{i}\mathbf{T}_{i+1}$ denote the coordinate transformation matrix from the coordinate system of the $(i+1)$st joint to that of the ith joint. The transformation matrix is given by

$$ {}^{i}\mathbf{T}_{i+1} = {}^{i}\,\mathbf{T}_{i'}\,{}^{i'}\mathbf{T}_{i+1}, \tag{1}$$

where

$$ {}^{i}\mathbf{T}_{i'} = \mathbf{Q}_i = \mathbf{Rot}_z(q_i), \tag{2}$$

$$ {}^{i'}\mathbf{T}_{i+1} = \mathbf{V}_i, \tag{3}$$

$$ \mathbf{Rot}_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

$$ \mathbf{Trans}\left(\begin{bmatrix} x & y & z \end{bmatrix}^t\right) = \begin{bmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5}$$

$$ q_i = s_i\, q'_i, \tag{6}$$

$s_i \in \{-1, +1\}$ is a sign parameters, q'_i is the ith joint value,

$$ \mathbf{V}_i = \mathbf{Rot}_z\,(\beta_i)\ \mathbf{R}_i\,\mathbf{Trans}(\begin{bmatrix} l_{i,x} & l_{i,y} & l_{i,z} \end{bmatrix}^t), \tag{7}$$

and

$$ \mathbf{R}_i = \begin{bmatrix} 1-\frac{b_{i,x}^2}{1+b_{i,z}} & \frac{-b_{i,x}b_{i,y}}{1+b_{i,z}} & b_{i,x} & 0 \\ \frac{-b_{i,x}b_{i,y}}{1+b_{i,z}} & 1-\frac{b_{i,y}^2}{1+b_{i,z}} & b_{i,y} & 0 \\ -b_{i,x} & -b_{i,y} & b_{i,z} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{8}$$

The rotation angle β_i and the translation parameter $l_{i,z}$ can be set to zero when a 4-DOF kinematic model is required [28]. Otherwise, they can also be treated as free parameters to offer a 6-DOF kinematic model.

According to [27], human joints have different DOFs, and every limb has different dynamic range and constraints. The DOFs and the terms of the movements of each body parts are listed in Table II. Joints having more than one DOFs are formulated as multiple revolute joints whose rotation axes have a common intersection point in the CPC model. If the pivot position between the ith joint and the $(i+1)$st joint are coincided, the translation vector in equation (7) can be set to zero.

Figure 3 shows the kinematic chain of an upper limb, where $\{W\}$ denotes the world coordinate system and and $\{\text{base}\}$ is the base (i.e., the torso) coordinate system. The shoulder and the elbow are formulated as three ($\mathbf{Q}_1$, $\mathbf{Q}_2$, and $\mathbf{Q}_3$) and two ($\mathbf{Q}_4$, and $\mathbf{Q}_5$) revolute joints, respectively. The coordinate systems of the accelerometers attached to the upper arm and the forearm are defined with two extra 6-DOF transformation matrices $\mathbf{V}_{acc1}$ and $\mathbf{V}_{acc2}$, respectively. Therefore, the coordinate transformation matrix from the accelerometers to the world coordinate system are respectively given by

$$ {}^{W}\mathbf{T}_{acc1} = {}^{W}\mathbf{T}_0\mathbf{V}_0\mathbf{Q}_1\mathbf{V}_1\mathbf{Q}_2\mathbf{V}_2\mathbf{Q}_3\mathbf{V}_{acc1}, \quad (9) $$

and

$$ {}^{W}\mathbf{T}_{acc1} = {}^{W}\mathbf{T}_0\mathbf{V}_0\mathbf{Q}_1\mathbf{V}_1\mathbf{Q}_2\mathbf{V}_2\mathbf{Q}_3\mathbf{V}_3\mathbf{Q}_4\mathbf{V}_4\mathbf{Q}_5\mathbf{V}_{acc2}. \quad (10) $$

To use the CPC kinematic model in video tracking, it is assumed that the monocular camera is calibrated such that its intrinsic parameters and its position and orientation with respect to $\{W\}$ are known. In order to fit the generic CPC model to different subjects, three translation parameters, i.e., $l_{0,x}$, $l_{0,y}$, and $l_{3,z}$, are treated as unknowns to be estimated in the posture reconstruction process. Notably, $l_{0,x}$ and $l_{0,y}$ determine the position of the shoulder joint while $l_{3,z}$ determines the length of the upper arm. The kinematic chain of a lower limb is defined in a similar way and, thus, the kinematic chains of the four limbs share the same world coordinate system and the same base coordinate system. Each limb has five joint angles and three translation parameters to be estimated. Additionally, the torso has six DOFs to be determined. Therefore, at each time instance t, the human posture estimation problem has 38-D unknown vector denoted as $\boldsymbol{x}_t$ to be estimated. Since searching an optimal solution in the 38-D space is intractable, we will assume that a reliable initial solution $\boldsymbol{x}_0$ is known and the posture reconstruction problem reduces to a state tracking problem which will be detailed in the next section.

C. Human Posture Tracking

The human posture tracking problem is an ill-posed problem which possesses high nonlinearity in both measurement functions and the state transition function and it also has multiple local optima. To handle the nonlinearity and the local optima, the particle filtering technique is adopted in this work. The particle filtering is capable of modeling nonlinear and multi-modal posterior distributions. The update of the posterior probability distribution of the state vector is governed by the following Chapman-Komogorov equation.

$$ p(\boldsymbol{x}_t|\boldsymbol{z}_t) \propto p(\boldsymbol{z}_t|\boldsymbol{x}_t)\int p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})p(\boldsymbol{x}_{t-1}|\boldsymbol{z}_{t-1})d\boldsymbol{x}_{t-1}, \quad (11) $$

where $\boldsymbol{z}_t$ is the measurement vector consisting of the silhouette image s of the subject and the tri-axis acceleration measurements of the nine accelerometers $\mathbf{a}_i \in \mathbb{R}^3$, i = 1, 2, ..., 9. For simplicity, we assume that the measurements are mutually independent. Therefore, the likelihood function $p(\boldsymbol{z}_t|\boldsymbol{x}_t)$ can be formulated as follows.

$$ p(\boldsymbol{z}_t|\boldsymbol{x}_t) = p(\mathrm{s}|\boldsymbol{x}_t)\prod_{i=1}^{9} p(\mathbf{a}_i|\boldsymbol{x}_t), \quad (12) $$

Assume that the measurement noises of the accelerometers are Gaussain, the likelihood of the accelerations is given by

$$ p(\mathbf{a}_i|\boldsymbol{x}) \propto \exp -\frac{\|\mathbf{a}_M - \mathbf{a}_i\|^2}{2\sigma_a^2}, \quad (13) $$

where $\mathbf{a}_M = \mathbf{g} + \mathbf{a}_b$ is composed of the acceleration of gravity $\mathbf{g}$ and the acceleration of body part motion $\mathbf{a}_b$. Both the two acceleration sources can be computed using the kinematic model. Returning to

TABLE II
TERMS OF MOVEMENTS

Body part	DOFs	Terms of movements
LUA/RUA	3	Abduction/Adduction Extension/Flextion Lateral rotation/Medial rotation
LFA/RFA	2	Extension/Flexion Supination/Pronation
LT/RT	3	Abduction/Adduction Extension/Flexion Lateral rotation/Medial rotation
LS/RS	2	Extension/Flexion Circumduction

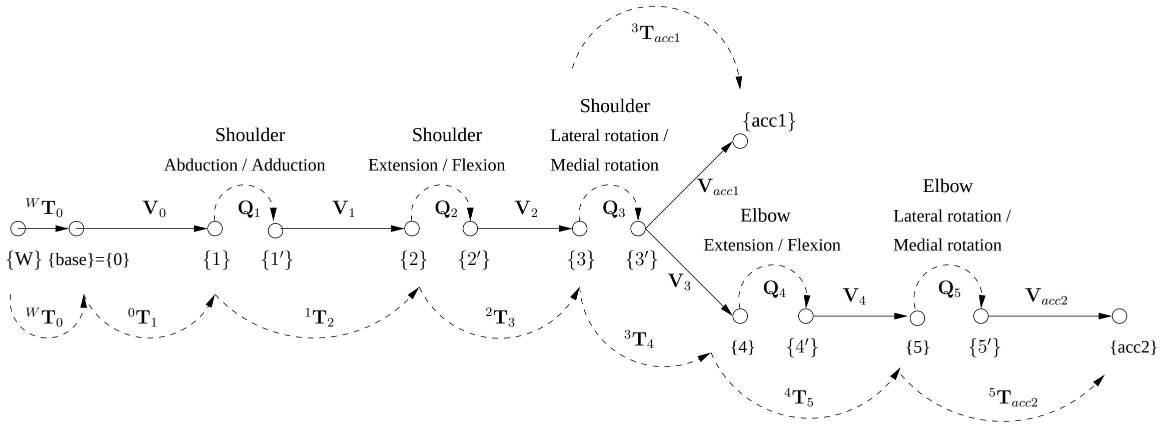


Fig. 3. The kinematic chain of an upper limb.

the kinematic chain example of Fig. 3, the acceleration of accelerometer acc1 due to body part motion can be computed using the second order derivative of the sensor position.

$$ {}^{acc1}\mathbf{a}_b = {}^{W}\mathbf{T}_{acc1}^{-1} \frac{d^2\left({}^{W}\mathbf{T}_{acc1}\mathbf{e}_0\right)}{dt^2}, \tag{14} $$

where $\mathbf{e}_0 = [0\ 0\ 0\ 1]^\top$. The acceleration of the gravity measured by acc1 can also be computed as follows.

$$ {}^{acc1}\mathbf{g} = -{}^{W}\mathbf{T}_{acc1}^{-1}\,[0\ 0\ 9.8\ 0]^\top, \tag{15} $$

where it is assumed that the acceleration of the gravity is in the opposite direction of the z-axis of $\{W\}$.

The likelihood function of the silhouette image can also be evaluated using the kinematic model. Let s_M denote the predicted silhouette of a body part computed by using the kinematic equations and the camera parameters. The likelihood of s is defined as follows.

$$ p(\mathrm{s}|\boldsymbol{x}_t) \propto \exp - \frac{\left|(\mathrm{s}\otimes \mathrm{s}_M)\cdot \mathrm{s}_M\right|^2 / \left|\mathrm{s}_M\right|^2}{2\sigma_s^2}, \tag{16} $$

where $|\cdot|$ denotes the number of non-zero pixels, and '$\otimes$' and '$\cdot$' are the logical XOR and the logical AND operations, respectively. The numerator of the exponent of (16) measures the normalized difference between the silhouettes of the measured and the predicted.

The particle filtering procedure are outlined in the following three steps.

Step 1. **Particles Sampling**
The discrete time propagation of the state density is derived from $\int p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) \cdot p(\boldsymbol{x}_{t-1}|\boldsymbol{z}_{t-1})d\boldsymbol{x}_{t-1}$. New particles, $\{\boldsymbol{x}_t^i, i = 1, 2, ..., N\}$, are sampled from previous weighted particles with the specified transition function.

Step 2. **Measurement and Particles Weighting**
The weight of particle i, denoted as π_t^i, is assigned as follows.

$$ \pi_t^i = k\cdot p(\boldsymbol{z}_t|\boldsymbol{x}_t = \boldsymbol{x}_t^i), \tag{17} $$

where k is a normalization constant such that $\sum_i \pi_t^i = 1$. The likelihood function $p(\boldsymbol{z}_t|\boldsymbol{x}_t = \boldsymbol{x}_t^i)$ is defined in equation (12).

Step 3. **State Estimating**
Each body part state $\boldsymbol{x}_t$ at each time step t can then be estimated by

$$ \boldsymbol{x}_t = \boldsymbol{x}_t^{i^*}, \tag{18} $$

where $i^* = \arg\max_i(\pi_t^i)$.

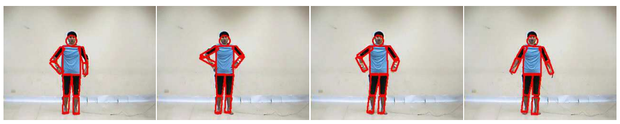
The body parts are tracked from the torso toward the end of the limbs.

IV. Experimental Results

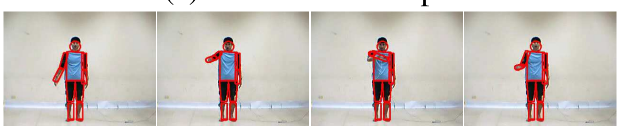
The camera used in the experiment is Logictech QuickcamPro which is calibrated using the method proposed in [29]. The accelerometers adopted in this work is Hitachi H34C. The video sequences used in the experiments are acquired at 10 frames per second with a resolution of 320 by 240 pixels. The accelerometer data is acquired synchronously.

Silhouettes of the target subject are segmented using a pre-trained background model.

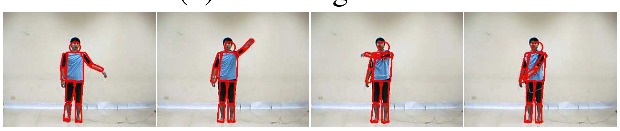
In the first experiment, we tested the propose method using simple actions such as hands on the hips, checking watch, circling, hands fold across the chest, and extension. The posture reconstruction results are shown in Fig. 4, where the wire-frame of the recovered human posture are overlayed on the input image. The results shows that our method successfully tracks the human posture in simple video sequences.



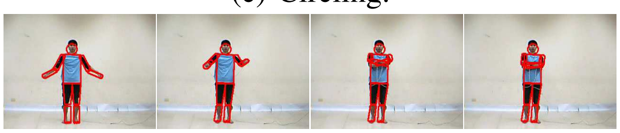
(a) Hands on the hips.



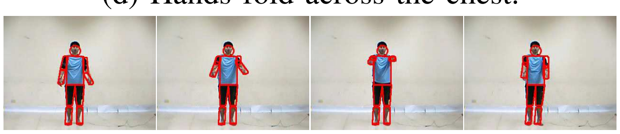
(b) Checking watch.



(c) Circling.



(d) Hands fold across the chest.



(e) Extension.

Fig. 4. Tracking results for simple actions.

In the second experiment, the proposed method is applied to track a fundamental Tai-Chi action which is termed *"ward off", "roll back", "press", and "push" (i.e.,* 掤、捋、擠、按*)*. Although this Tai-Chi action possesses high order self-occlusion, the tracking results shown in Fig. 5 are satisfactory. In order to show the importance of using the accelerometers, we also track this Tai-Chi action without using any inertial sensors for comparison. As shown in Fig. 6, the tracking results degraded severely because the monocular silhouette sequences is ambiguous to recover the 3-D human posture. The results demonstrate that inertial sensors can improve 3-D tracking results.

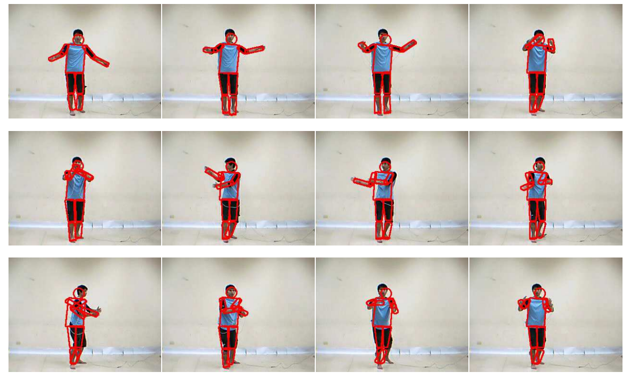


Fig. 5. Tracking results of "ward off", "roll back", "press", and "push"

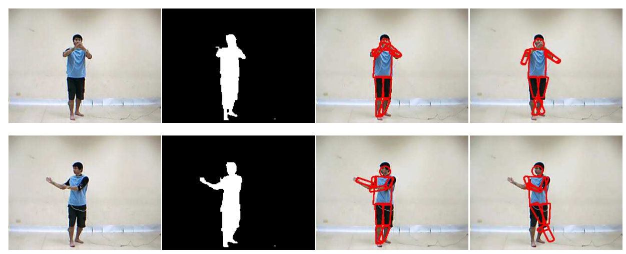


Fig. 6. The observed image, segmented silhouette, tracking result with the proposed method, and tracking result without using inertial sensors are shown from left to right.

V. Conclusions and Future Works

In this work, a human posture tracking method is proposed which uses a monocular camera and multiple accelerometers attached to different body parts of a subject. To simplify the 3-D posture tracking problem in a monocular video sequence, a generic CPC kinematic model [19] of the human body is created. Twelve CPC parameters are adapted according to the lengths of the upper limbs and the joint positions of the shoulders and the thighs so that the generic CPC model can be used to track different subjects. The particle filter technique is used to track the body parts in a top-down sequence from the torso toward the ends of the limbs. The proposed system is cost-effective which can be used to track human motion for assisting the learning of Tai-Chi actions. Real experiments have been conducted, and the results show that the performance of the proposed method is satisfactory.

Future work will focus on constructing a more complete kinematic model to describe the parameters of body parts and incorporating information provided by other sensors, such as rate-gyros, to construct a more reliable likelihood function. We will also impose some human anthropometric constraints to avoid generating impossible human postures. Furthermore, it is desired to carry out quantitative evaluation of our method.

ACKNOWLEDGEMENTS

This work is supported in part by the National Science Council of Taiwan under grant NSC95-2221-E260-028-MY3.

REFERENCES

[1] P. T. Chua, R. Crivella, B. Daly, N. Hu, R. Schaaf, D. Ventura, T. Camill, J. Hodgins, and R. Pausch. Mastermotion: Full body wireless virtual reality for tai chi. In *SIGGRAPH '02: ACM SIGGRAPH 2002 conference abstracts and applications*, pages 214–214, New York, NY, USA, 2002. ACM.

[2] K. Meyer, H. L. Applewhite, and F. A. Biocca. A survey of position trackers. *Presence*, 1(2): 173–200, 1992.

[3] H. Dejnabadi, B. M. Jolles, and K. Aminian. A new approach to accurate measurement of uniaxial joint angles based on a combination of accelerometers and gyroscopes. *IEEE Transactions on Biomedical Engineering*, 52 (8):1478–1484, 2005.

[4] H. Dejnabadi, B. M. Jolles, E. Casanova, P. Fua, and K. Aminian. Estimation and visualization of sagittal kinematics of lower limbs orientation using body-fixed sensors. *IEEE Transactions on Biomedical Engineering*, 53 (7):1385–1393, 2006.

[5] Q. Delamarre and O. Faugeras. 3d articulated models and multi-view tracking with silhouettes. In *International Conference on Computer Vision*, volume 2, pages 716–721, 1999.

[6] I. Kakadiaris and D. Metaxas. Model-based estimation of 3d human motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1453–1459, 2000.

[7] G. K. M. Cheung, S. Baker, and T. Kanade. Shape-from-silhouette of articulated objects and its use for human body kinematics estimation and motion capture. In *Computer Vision and Pattern Recognition*, volume 1, pages 77–84, 2003.

[8] R. Kehl, M. Bray, and L. V. Gool. Full body tracking from multiple views using stochastic sampling. In *Computer Vision and Pattern Recognition*, volume 2, pages 129–136, 2005.

[9] B. Michoud, E. Guillou, and S. Bouakaz. Shape from silhouette: Towards a solution for partial visibility problem. In *Eurographics Short Papers Preceedings*, 2006.

[10] B. Michoud, E. Guillou, H. Brice no, and S. Bouakaz. Real-time marker-free motion capture from multiple cameras. In *International Conference on Computer Vision*, pages 1–7, 2007.

[11] T. Liu, Y. Inoue, and K. Shibata. A wearable sensor system for human motion analysis and humanoid robot control. In *IEEE International Conference on Robotics and Biomimetics*, pages 43–48, 2006.

[12] H. J. Luinge and P. H. Veltink. Inclination measurement of human movement using a 3-d accelerometer with autocalibration. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 12(1):112–121, 2004.

[13] D. T. W. Fong, J. C. W. Wong, A. H. F. Lam, R. H. W. Lam, and W. J. Li. A wireless motion sensing system using adxl mems accelerometers for sports science applications. In *Proceedings of Fifth World Congress on Intelligent Control and Automation*, volume 6, pages 5635–5640, 2004.

[14] H. Junker, O. Amft, P. Lukowicz, and G. Tröster. Gesture spotting with body-worn inertial sensors to detect user activities. *Pattern Recognition*, 41(6):2010–2024, 2008.

[15] R. Zhu and Z. Zhou. A real-time articulated human motion tracking using tri-axis inertial/magnetic sensors package. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 12(2):295–302, 2004.

[16] J. A. Ward, P. Lukowicz, G. Tröster, and

T. Starner. Activity recognition of assembly tasks using body-worn microphones and accelerometers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(10): 1553–1567, 2006.

[17] A. Hauck, S. Lanser, and C. Zierl. Hierarchical recognition of articulated objects from single perspective views. In *Computer Vision and Pattern Recognition*, pages 870–876, 1997.

[18] M. Isard and A. Blake. Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29 (1):5–28, 1998.

[19] H. Zhuang, Z. S. Roth, and F. Hamano. A complete and parametrically continuous kinematic model for robot manipulators. *IEEE Transactions on Robotics and Automation*, 8 (4):451–463, 1992.

[20] H. H. Chi, J. Song, and G. Corbin. "killer app" of wearable computing: wireless force sensing body protectors for martial arts. In *UIST*, pages 277–285, 2004.

[21] M. Takahata, K. Shiraki, Y. Sakane, and Y. Takebayashi. Sound feedback for powerful karate training. In *NIME*, pages 13–18, 2004.

[22] K. Kunze, M. Barry, E. A. Heinz, P. Lukowicz, D. Majoe, and J. Gutknecht. Towards recognizing tai chi - an initial experiment using wearable sensors. In *Proceedings of IFAWC*, pages 380–389, 2006.

[23] D. Spelmezan and J. Borchers. Real-time snowboard training system. In *CHI Extended Abstracts*, pages 3327–3332, 2008.

[24] D. Y. Kwon and M. H. Gross. Combining body sensors and visual sensors for motion training. In *Advances in Computer Entertainment Technology*, pages 94–101, 2005.

[25] E. Schoonderwaldt, N. H. Rasamimanana, and F. Bevilacqua. Combining accelerometer and video camera: Reconstruction of bow velocity profiles. In *Proceedings of the 2006 International Conference on New Interfaces for Musical Expression*, pages 200–203, 2006.

[26] R. E. Mayagoitia, A. V. Neneb, and P. H. Veltink. Accelerometer and rate gyroscope measurement of kinematics: an inexpensive alternative to optical motion analysis systems. *Journal of Biomechanics*, 35:537–542, 2002.

[27] K. L. Moore and A. F. Dalley. *Clinically Oriented Anatomy*. Lippincott Williams and Wilkins, 4 edition edition, April 1999.

[28] S. W. Shih, Y. P. Hung, and W. S. Lin. Kinematicparameter identification of an binocular head using stereo vision. Technical Report TR-94-006, Institute of Information Science, Academia Sinica, Nankang, Taipei, Taiwan, 1994.

[29] Z. Zhang. Flexible camera claibration by viewing a plane from unknown orientations. In *International Conference on Computer Vision*, pages 666–673, 1999.