CrossMark

# Smart motion reconstruction system for golf swing: a DBN model based transportable, non-intrusive and inexpensive golf swing capture and reconstruction system

Dongyue Lv[1,2] · Zhipei Huang[1] · Lixin Sun[1] ·
Nenghai Yu[2] · Jiankang Wu[1]

**Abstract** In the past decade, golf has stimulated people's great interest and the number of golf players has increased significantly. Therefore, how to train a golfer to make a perfect swing has attracted extensive research attentions. Among these researches, the most important step is to capture and reconstruct the swing movement in a transportable and non-intrusive way. Restricted by the development of present depth imaging devices, the initial captured swing movement may not be acceptable due to occlusions and mixing up of body parts. In this paper, to restore motion information from self-occlusion and reconstruct 3D golf swing from low resolution data, a Dynamic Bayesian Network (DBN) model based golf swing reconstruction algorithm is proposed to increase the capture accuracy considering the spatial and temporal similarities of swing between different golfers. A Smart Motion Reconstruction system for Golf swing, SMRG, is presented based on the DBN model with a popular depth imaging device, Kinect, as capturing device. Experimental results have proved that the proposed system can achieve comparable reconstruction accuracy to the commercial optical motion caption (OMocap) system and better performance than state of art modification algorithms using depth information.

✉ Zhipei Huang
  zhphuang@ucas.ac.cn

[1]  University of Chinese Academy of Sciences, Beijing, China

[2]  Institution of Electronics, Chinese Academy of Sciences, Beijing, China

 Springer

# 1 Introduction

Golf is one of the most developed and popular sports. The number of active golf players is estimated at 50 million and golf course has become the fastest growing development all over the world. However, how to train a golfer to make a proper and accurate golf swing movement is critical to the success of the golf training course [17]. Previously, the training is usually guided by professional golfers, which involves extensive human recourse. In recent years, using standard swing procedures for golf training has become a tendency. A standard (perfect) swing is considered containing many important parameters, such as club head trace [4, 8, 15, 30], swing plane [11], hub path [21], leg movement [9] and even wrist angle [3, 29]. However, to extract these parameters, the key is to accurately capture the swing movement in 3D space.

For this purpose, two types of motion capture (Mocap) systems are mainly applied in recent years [35]: optical Mocap (OMocap) system [7, 16, 20, 32] and wearable micro-sensors based Mocap system (MMocap) [1, 2, 5, 13, 25, 31]. The OMocap systems require attaching reflective markers on golfer's body segment. The positions of the markers are obtained via multiple fixed high-speed cameras around the golfer. The MMocap systems also require placing micro-sensors on golfer's body segments for motion reconstruction. The motion of segments is then acquired by calculating sensor information. Although these two types of Mocap systems have been used successfully in many swing applications, the OMocap system restricts the application environment while the MMocap system requires golfer to wear some extra devices. These limitations are intrusive and make the golfer uncomfortable during the swing movement-not to mention the expensive cost and complicated system formation.

Recently, the development of depth imaging devices shows a tendency of transportable and non-intrusive way to capture golf swing. With low cost and convenient placement, these devices have been increasingly applied to 3D motion capture researches [6, 10, 14, 22, 24, 27]. However, restricted to present technology, the resolution of depth images is barely acceptable and the main challenge is its poor performance due to self-occlusions and mixing up of body parts [27]. To deal with these two issues and make it applicable for the golf swing analysis, some ad-hoc solutions to improve the reconstruction accuracy have been proposed so far. For example, Zhang et al. [33, 34] proposed to capture the 3D skeleton coordination of a golfer while performing swing, and then used serial correlation model to score and recognize grade of golf swing. Lin et al. [12] captured the swing motion and automatically identify the 6 common mistakes on swing motion. Shen et al. [26] tried to tackle the occlusion problem and presented an exemplar-based method to learn and correct the initially estimated poses. Although these methods have made some progress to improve the wrongly recognized swing posture, the challenge is still not very well solved yet due to the ignorance of motion similarity contained in the swing dynamics.

In this paper, a Dynamic Bayesian Network (DBN) model based golf swing reconstruction algorithm is proposed to improve the capture performance of depth imaging devices, and a golf swing capture and reconstruction system (SMRG) is built based on this algorithm. To improve the performance against self-occlusion and gain accurate joints positions from low resolution data, the algorithm integrates the spatial and temporal similarities among joints and their movement dynamics into the DBN model. Although not clearly mentioned in the above studies, their algorithms and the idea "standard (perfect) swing" in golf training have all avowed the similarities between swings. Moreover, the similarities exist not only in the same person, but also between different people. They can be divided into two parts: 1) spatial similarity: the relative movements among joints during swing are similar if the swings are

performed without external interference; 2) temporal similarity: the swing periods between different golfers maybe different, but the normalized partitions of the four stages (stance, backswing, downswing and follow-through) in their swing are similar. These similarities, if applied properly, can also be used to further improve the accuracy of motion information against self-occlusion and reconstruct 3D golf swing from low resolution depth image sequences. In the proposed DBN model, the spatial similarity is integrated by applying the same model structure (order number and parameters) and the temporal similarity is integrated by normalizing the real swing periods. The model is trained using actual golf swing motion data from different golfers captured by an OMocap system MAT-T [18]. The initial joint positions generated from low resolution depth images are optimized and the reconstruction of golf swing is performed using these optimized joints positions. We have compared our position outputs with a commercial OMocap system MAT-T and two typical depth imaging device based motion reconstruction systems proposed in [27] and [26]. The good performance results have shown that our system can achieve comparable tracking accuracy to the MAT-T system and improves the depth imaging devices' outputs much more than the other two systems.

The rest of the paper is organized as follows: Section 2 describes the DBN model based SMRG system. Experimental results are given in Section 3. Finally, conclusions and future work are provided in Section 4.

### 1.1 Ethics statement

We did not seek approval in this case. The reasons are:

1) Golf swing cannot do any harm to the participants.
2) The swing data were analyzed anonymously.

We have gained personal authorizations of the five participants to use their swing data in our study. The five participants all signed formal informed consents before the swings.

## 2 SMRG: the DBN model based swing reconstruction system

### 2.1 System framework

The system contains three parts: motion data acquisition, motion data processing and swing reconstruction. The framework of our system is shown in Fig. 1. As a typical depth imaging device, the Microsoft Kinect has gained full software support and shown some potential to overcome the pitfalls of the traditional Mocap systems [28]. In the first part, a real world swing is captured by Kinect and represented as RGB-D images. The images are then transferred into
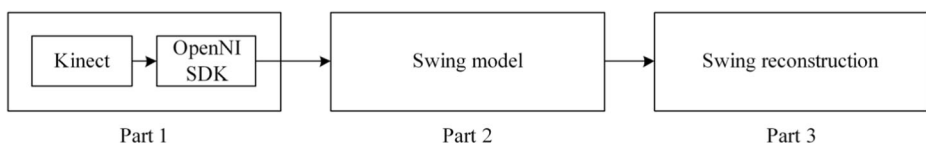


**Fig. 1** System framework

motion data using OpenNI SDK [23]. The second part modifies the motion data with the proposed DBN model. The third part reconstructs the swing using the modified motion data.

## 2.2 Motion data acquisition

We apply joints positions as motion data to represent golfer's movement. The Kinect generates RGB-D images of golfer and swing scene. In practice, the Kinect is put in front of the golfer with a distance about 2.5 m with a height about 1 m from the ground to make sure the whole swing motion can be captured. The images are then transmitted to computer to acquire motion data. To increase the mobilization of our system, we modify the data transmission mode into wireless. The OpenNI SDK is applied to transfer the Kinect captured images into usable motion data. This SDK is based on Windows Kinect SDK (OpenNI 2.0+) and the joint position observations are provided by its skeleton functions. These functions can generate preliminary joint positions from the RGB-D images.

## 2.3 Motion data processing: the DBN model

The initial joint positions generated from depth images cannot be directly used since the resolution of the depth images is quite low (restricted by present technology) and the problem of self-occlusion is not well solved. A five chains full-body DBN model is proposed to restore accurate joint positions from initial ones. In the model, both the spatial and temporal similarities are integrated. The spatial similarity is integrated by applying the same model structure (order number and parameters) and the temporal similarity is integrated by normalizing the real swing periods.

### 2.3.1 DBN model structure

We consider the human body as a five chains skeleton model which is shown in Fig. 2 (left). Each chain is a set of segments connected by joints. The segments in a chain are considered as rigid bodies. To reconstruct the golf swing accurately, the exact positions of the key joints should be acquired.

In the above golfer model, 15 key joints include: head, neck, torso, left and right shoulders, left and right elbows, left and right hands, left and right hips, left and right knees and left and right feet. These 15 key joints construct a hierarchical structure of a golfer containing the five serial chains, which is shown in Fig. 2 (right). In the hierarchical structure, the torso joint is considered as the root of all the chains. To construct our DBN model, the positions of joints are focused on one arbitrary chain. It can be derived similarly in other four chains by using the DBN model of the chosen chain.

In the DBN model, 5 symbols are used to represent the states and observations. In the description below, JCS is a coordinate system with one moving joint as the original point (for example, the right hand's relative position is measured in the coordinate system with the right elbow as the original point), and GCS is the coordinate system with one fixed point as the original point.

$s_t^i$    The relative position of the ith joint in chain in its parent joint's coordinate system (JCS) in time t.

$X_t^i$    The absolute position of the ith joint in chain in global coordinate system (GCS) in time t.

$Y_t^i$    The position observation (the coordinates given by Kinect and OpenNI SDK) of the ith joint in chain in GCS in time t.
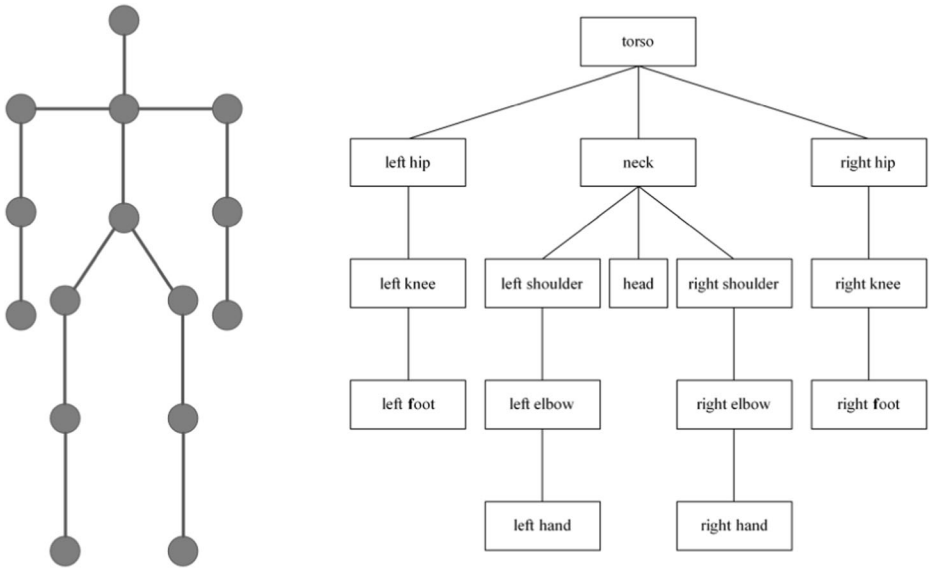
**Fig. 2** Human skeleton model (left) and hierarchical structure (right)

$n$    The order number describing the dynamics of $s_t^i$ in the whole swing.

$m$    The number of joints in the chain.

The exact structure of the DBN model of one chain is shown in Fig. 3. The static structure of joint chain and the dynamic structure of one joint are shown respectively. During model construction, the dynamics of the relative position of one joint could be first-order or multi-order Markov. The precise order number "n" is determined through practice.

The purpose of the DBN model is to estimate more precise positions of the joints in GCS during the whole swing, i.e., find the most probable position of $X_t^i$ by calculating the posterior probability $\Pr(X_t^i|Y_{1:t}^{0:i})$. Based on the structure of the proposed DBN model, a preliminary iteration sequence is raised to solve the estimation problem:

$$X_1^0, X_1^1, \ldots, X_1^m, X_2^0, \ldots, X_2^m, \ldots, X_t^0, \ldots, X_t^m \tag{1}$$

This means reconstructing the swing of the golfer accurately, every joint's position in GCS from every frame should be acquired precisely. According to (1), the position estimation results should be acquired sequentially, following the chain order in one frame, and in the next frame, the same order should be followed. The position estimation $X_t^i$ is related to its parent $X_t^{i-1}$ in the same frame and its relative position $s_{t-1}^i$ in previous frame. A modified iteration sequence is proposed to estimate every joint position in chain in time t:

$$s_{t-1}^0, X_t^0, \ldots s_{t-1}^m, X_t^m \tag{2}$$

The inference of our model needs three key elements: Spatial relationship $\Pr(X_t^i|X_t^{i-1}, s_t^i)$, temporal relationship $\Pr(s_t^i|s_{t-1:t-n}^i)$ and likelihood $\Pr(Y_t^i|X_t^i)$. For defining and training simplicity, all these three elements are assumed to follow normal distribution. After definition, the parameters in these three elements can be learned before reconstruction by training previous motion data gained from the same golfer.
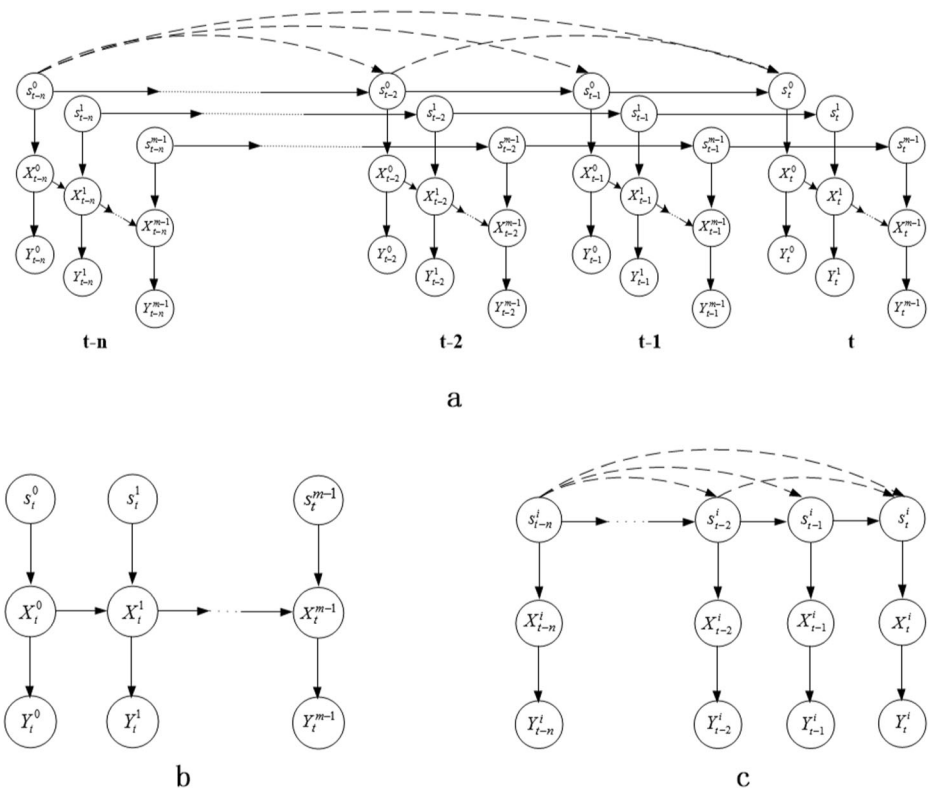
**Fig. 3** The DBN model of one arbitrary chain of human model, **a** is the whole structure, **b** is the static structure and **c** is the dynamic structure

**Spatial relationship** According to the human skeleton model, the joints positions are the only elements been discussed. Therefore the states $s_t^i$ and $X_t^i$ are both joint positions. The difference between the two kinds of states is that $s_t^i$ is the position of one joint in its parent's JCS, while $X_t^i$ is the position in GCS. Apparently, $X_t^i$ can be gained by integrating $s_t^i$ and $X_t^{i-1}$, according to the chain structure. The spatial relationship between parent and child joints can be written as:

$$\Pr\left(X_t^i \middle| X_t^{i-1}, s_t^i\right) = N\left(X_t^{i-1} + s_t^i, Q_i\right) \tag{3}$$

In (3), $Q_i$ is the process noise from previous training sets and their ground truths of one golfer.

**Likelihood** The observation $Y_t^i$ is generated from state $X_t^i$. Since they are all measured in GCS and only have measurement errors, the likelihood can be written as:

$$\Pr\left(Y_t^i \middle| X_t^i\right) = N\left(X_t^i, R_i\right) \tag{4}$$

In (4), $R_i$ is the measurement noise from the OpenNI SDK outputs.

**Temporal relationship** For different golfers, the motion of each person's joints should be unique but repeatable if the swing is performed again by the same person. This assumption is taken into consideration to train different motion models of each joint for different golfers. The motion of joints is only considered in their parents' JCS to eliminate the affection of their parent joints and acquire 'pure' motion models. Each joint can acquire a motion model by training previous motion data. Following the normal distribution, the temporal relationship of one joint can be described as:

$$\Pr\left(s_t^i \middle| s_{t-1:t-n}^i\right) = N\left(A_{t-1}^i \left[s_{t-1}^i, \ldots, s_{t-n}^i\right]^T, \Sigma_{t-1}^i\right) \tag{5}$$

In (5), the matrix $A_{t-1}^i$ and $\Sigma_{t-1}^i$ contain the parameters which need to be trained to get from $s_{t-1:t-n}^i$ to $s_t^i$.

**Order number selection** In practice, 5 golfers were participated in our experiment. Each of them repeated 6 swings. Since the proposed system is based on a multi-order DBN model, a proper order number should be determined beforehand. To evaluate the model performance on all the golfers, the mean value of all golfers' msJE [26] is applied and shown in Fig. 4. The order number rises from 1 to 8 since every training we have 24 training data. As can be seen in Fig. 4, the msJE drops first and rises when the order number is above 5. For the purpose of economize training and calculating consumption, apparently, the order number of our model is chosen as 5.

### 2.3.2 Swing period normalization

A whole swing contains four stages: stance, backswing, downswing and follow-through. We define "swing period" in our system from the beginning of backswing to a special posture in follow-through which is shown in Fig. 5. This period is the most activity during the whole swing. Due to different wrist force, gender and other factors, the swing periods of different golfers are mostly different. Normally to reconstruct swings, models should be built and
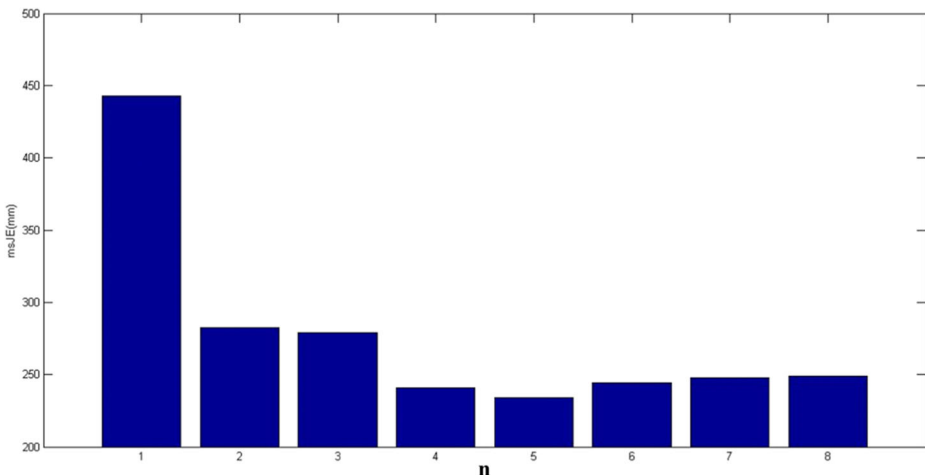


**Fig. 4** The variation of msJE with the order of model rises

Fig. 5 The special posture to end
swing period



trained for different golfers since the different swing periods. This limits the system generality. Sometimes one does not want to know the exact swing period; in that case, the only thing considered is whether the swing is perfect or not. So the swing period could be normalized in our system to expand the universality for different golfers.

We analyzed the motion of some key joints (hands, elbows and shoulders) from different golfers, which is shown in Fig. 6. It turns out that they have similar normalized motion
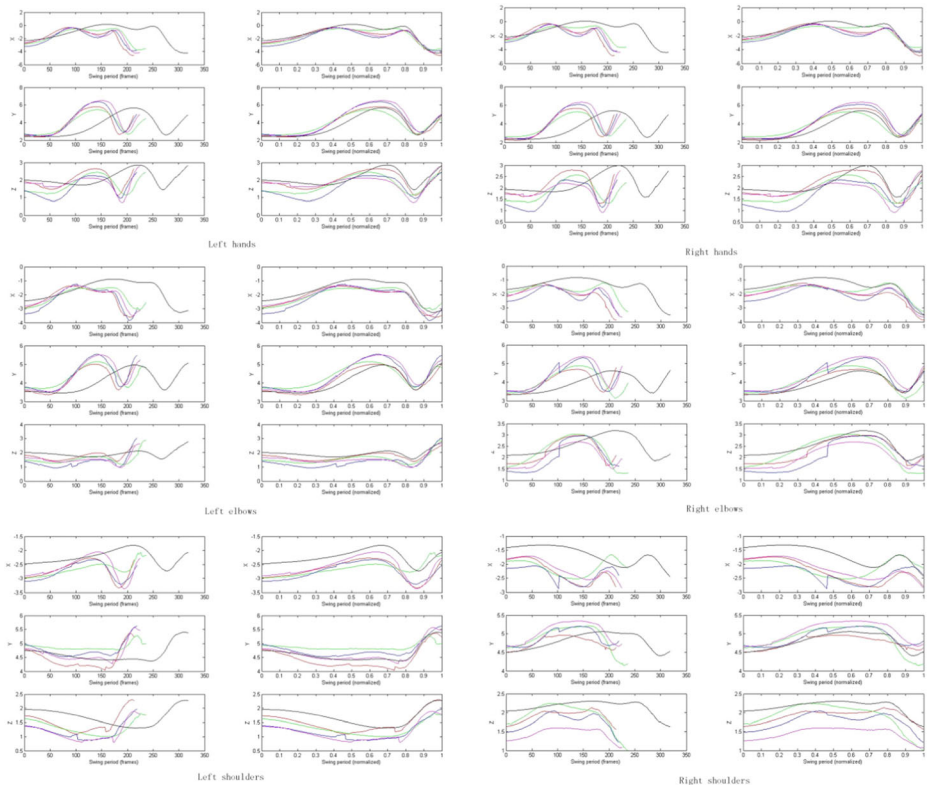


Fig. 6 The motion patterns of 6 key joints of 5 golfers with and without normalization

patterns. This similarity is the basis to perform normalization for different golfers. Moreover, the similarity is even the basis to build the DBN model, since only similar motion patterns could be well modeled in most cases. During our model construction, the discrete Kinect captured motion data will be normalized to a constant frame number. We counted the frame numbers Kinect captures when every golfer performs swing. The maximum frame number $N_{max}$ is chosen to be the normalized frame number. The piecewise linear interpolation is applied to make sure every swing period is normalized to $N_{max}$ in our system.

### 2.4 Swing reconstruction

The modified joints positions are applied to reconstruct swing in 3D space. In our system, golfers are drawn as skeleton models shown in Fig. 2 (left) with 4 angles of view. Although not aesthetically beautiful, the skeleton model can provide enough information to a golfer for training. An example frame of skeleton model based swing is shown in Fig. 7.

## 3 Experiments and discussion

### 3.1 Setup

The MAT-T system and the RFR algorithm [26] are used to evaluate the reconstruction performance of the SMRG system. The placement of the 6 cameras and the application environment of the MAT-T system are shown in Fig. 8.

5 golfers (4 males and 1 female) were participated in our experiment. Each of them repeated 6 swings (including stance, backswing, downswing and follow-through). During the experiment, all 5 golfers' swing data are used as testing set by turns, i.e., 4 golfers' swing data are used as training set while the other is testing set, then the testing set turns to training set and chose another golfer's data as testing set. This ensures every swing of every golfer can have a corresponding reconstruction result. The two systems (i.e., 6 cameras MAT-T system and Kinect) capture the swing synchronously. The sampling rate of the cameras from MAT-T system is 180Hz, while the Kinect is 30Hz.

**Fig. 7** An example frame of skeleton model

45 degree view
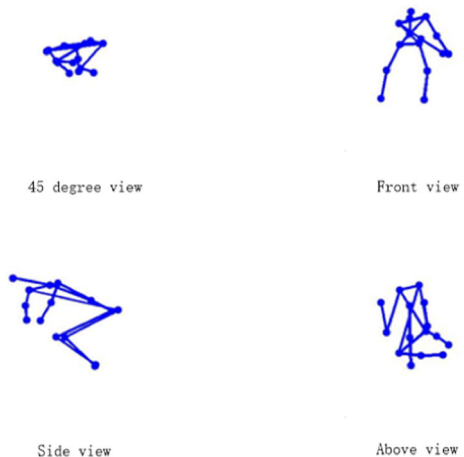
Front view

Side view

Above view

**Fig. 8** The application environment of the MAT-T system



After the training step, our system can perform an online modification with about 18 frames per second. To illustrate the feasibility of our system, the joints and body segments which have the most severe movements or rotations were chosen for comparison. In our implementation, we chose the hands positions to evaluate the performance. The shoulder width (SW) and arms lengths (left upper arm (LU), left lower arm (LL), right upper arm (RU) and right lower arm (RL)) were also taken into consideration.

The length of above 5 segments can be acquired indirectly by calculating the segment length between two relational joints. Comparing with outputs of MAT-T system, the error ratio $e$ between our outputs and MAT-T system outputs is used as a criterion:

$$e = \frac{\left|L - \tilde{L}\right|}{L} \tag{6}$$

In (6), $L$ is the segment length from MAT-T system, while $\tilde{L}$ is the one from our system.

Similar to the segment length comparison, the joint position difference between two kinds of outputs is applied. To evaluate the difference, as mentioned in [26], the mean value of the sum of joint errors (msJE) is calculated.

$$msJE = \frac{\sum\limits_{t=1}^{T} \sum\limits_{i=1}^{N} \left|X_t^i - \tilde{X}_t^i\right|}{T} \tag{7}$$

In (7), $X_t^i$ is the position of joint i in time t from MAT-T, and $\tilde{X_t^i}$ is from our system. T represents the swing period, which equals the normalized frame number $N_{max}$ Kinect captured during the whole swing.

As mentioned above, the msJE is applied as criterion. Because the number of the selected joints is different in three systems (15 in ours and 20 in the other two), a more precise value, the mean value of msJE (mmsJE) is also applied.

$$mmsJE = \frac{\sum\limits_{t=1}^{T} \sum\limits_{i=1}^{N} \left|X_t^i - \tilde{X}_t^i\right|}{T * N} \tag{8}$$

In (8), $N$ is the number of joints used to represent golfer, in our system $N=15$ while in the other two systems $N=20$.

## 3.2 Comparison with the MAT-T

We randomly chose one from each golfer's reconstruction results to perform the comparison. The hand position outputs of 5 golfers of our system and the MAT-T system are shown in Fig. 9, respectively. Because the joint moves in 3D space, the three components (x, y and depth) of positions are shown respectively.

In Fig. 9, the outputs of two kinds of results only have slightly difference. In general, our algorithm's performance is comparable with MAT-T in acquiring joints' positions. However, the cost of a Kinect is much cheaper than any existing commercial OMocap, including the MAT-T system. The main difference of two kinds of outputs is in depth component. The difference is a bit notable in follow-through stage. Although the DBN model improves the original Kinect outputs, the original hands' depth outputs are very unstable when severe occlusion and rapid movement occurs. These outputs will definitely affect the performance of the model.

The comparison results of 5 mentioned segments are listed in Table 1. The results are all mean values through the whole swing procedure.

Compared with the MAT-T system, the outputs of our system do have certain errors. Moreover, some segment's errors are significant (larger than 10 %). However, these errors
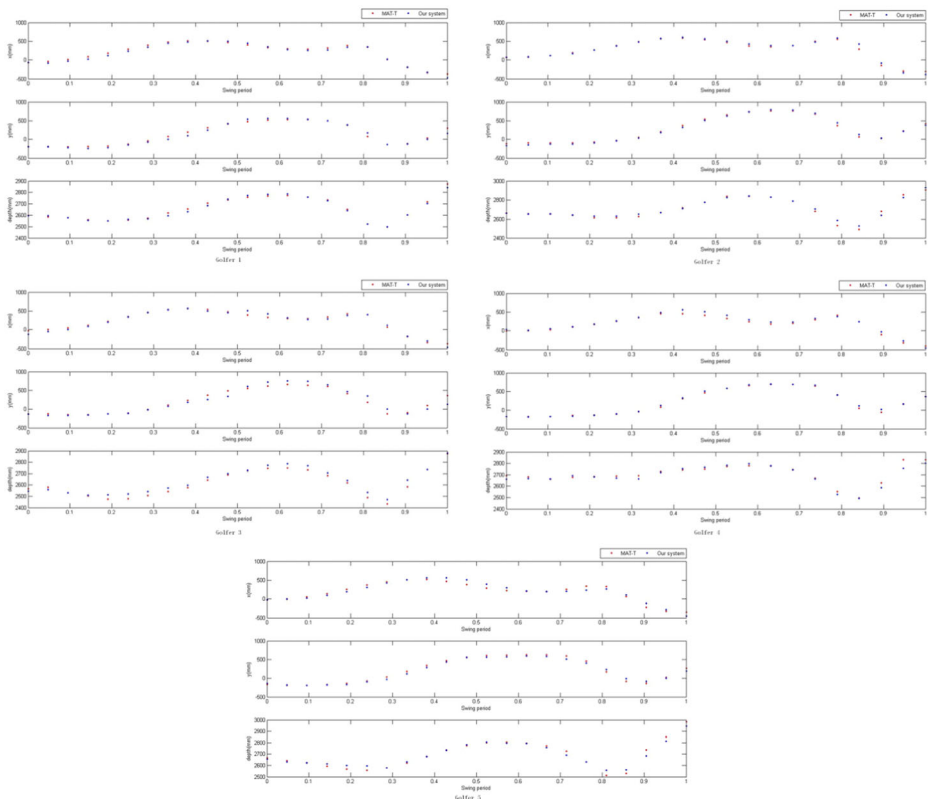


**Fig. 9** Comparison of hand position outputs

**Table 1** Comparison of segments lengths

|  | Segment | MAT-T (mm) | SMRG (mm) | e (%) |
|---|---|---|---|---|
| Golfer 1 | SW | 235.4584 | 265.0521 | 12.57 |
|  | LU | 250.2186 | 228.1173 | 8.83 |
|  | LL | 234.1757 | 247.2712 | 5.59 |
|  | RU | 250.2189 | 233.6072 | 6.64 |
|  | RL | 234.1752 | 252.2712 | 7.73 |
| Golfer 2 | SW | 246.8018 | 232.9003 | 5.63 |
|  | LU | 259.7177 | 248.6545 | 4.26 |
|  | LL | 243.0660 | 248.6550 | 2.30 |
|  | RU | 259.7176 | 274.4763 | 5.68 |
|  | RL | 243.0651 | 254.4766 | 4.69 |
| Golfer 3 | SW | 269.0449 | 305.3614 | 13.50 |
|  | LU | 303.9334 | 279.8513 | 7.92 |
|  | LL | 284.4474 | 272.6594 | 4.14 |
|  | RU | 303.9336 | 273.9884 | 9.85 |
|  | RL | 284.4457 | 302.1223 | 6.21 |
| Golfer 4 | SW | 275.8920 | 283.7762 | 2.86 |
|  | LU | 291.7776 | 279.8572 | 4.09 |
|  | LL | 273.0707 | 284.8576 | 4.31 |
|  | RU | 291.7773 | 281.4743 | 3.53 |
|  | RL | 273.0697 | 281.4735 | 3.08 |
| Golfer 5 | SW | 274.3791 | 281.7759 | 2.70 |
|  | LU | 297.5267 | 289.9451 | 2.54 |
|  | LL | 278.4502 | 259.1230 | 6.94 |
|  | RU | 297.5262 | 291.8557 | 1.91 |
|  | RL | 278.4491 | 275.6611 | 1.00 |

are not always mean bad performance. In the whole swing procedure, the skins of shoulders move more severely than any other body segments, because the shoulder joints rotate the most rapidly. Markers may change their position relative to their corresponding body segments due to skin movement [19]. If the shoulder markers change positions, the outputs of the MAT-T system will fluctuate, this may cause the difference. To demonstrate this, the SWs of the five golfers are shown in Fig. 10. The results are gained by our system, the MAT-T system and original Kinect output respectively.

In normal condition, the SW of each golfer should keep constant during swing, because the body segments are considered rigidly. In Fig. 10, the SWs of the MAT-T system are not constant. The main reason is that the shoulder markers have changed their position due to severe skin movements. The original Kinect outputs vary more severely due to the occlusion of shoulder joints. With the help of the DBN model, the outputs of our system show more robustness than the other two kinds of results.

There are some differences between the outputs of our system and the MAT-T system, but the feasibility of our algorithm is unquestionable. With the development of better capture device (better than Kinect but not as expensive as the MAT-T), it is no doubt that the performance will be better.
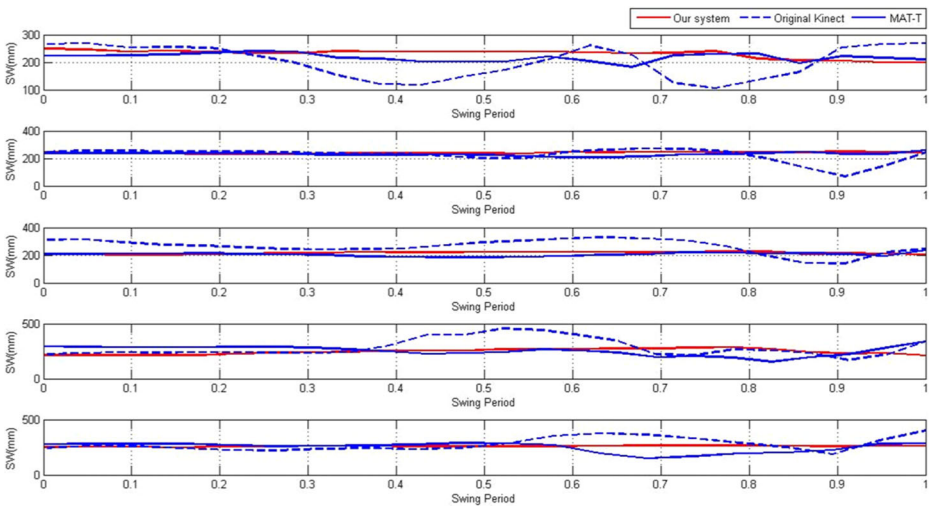
**Fig. 10** The shoulder widths of five golfers during whole swing

## 3.3 Comparison with other depth imaging device based reconstruction systems

In order to prove the superiority of our system to the state of art Kinect based motion reconstruction systems, the RFR system in [26] and the system mentioned in [27] are chosen. The reasons why these two systems are chosen as our opponents are a) though the rotation of golfer is limited, the test motion of RFR is golf swing and b) the test motions in [27] include self-occluded motions and golf swing is undoubtedly a self-occluded motion.

The two groups of values are listed in Table 2. The comparison shows that all the three systems can improve the Kinect performance, but our system performs much better.

By comparing the outputs of our system, the MAT-T system and two other Kinect-based systems, it can be found that our system works well with high robustness against severe joint occlusion and improves the Kinect performance. The outputs are comparable with the MAT-T system which illustrates the feasibility and effectiveness of the proposed algorithm and the system. The comparisons between our system and the two Kinect-based systems show that the mmsJEs of our outputs are much lower than the ones of the two other systems. Moreover, our

**Table 2** Comparison of msJE and mmsJE

|                        | msJE (mm) | mmsJE (mm)          |
| ---------------------- | --------- | ------------------- |
| Original Kinect output | 2071.6    | 103.58 (20 joints)  |
| The RFR system         | 1575.7    | 78.79 (20 joints)   |
| The system in [27]     | 1452.0    | 72.60 (20 joints)   |
| SMRG (Golfer 1)        | 245.71    | 16.38 (15 joints)   |
| SMRG (Golfer 2)        | 227.15    | 15.14 (15 joints)   |
| SMRG (Golfer 3)        | 291.41    | 19.43 (15 joints)   |
| SMRG (Golfer 4)        | 139.36    | 9.29 (15 joints)    |
| SMRG (Golfer 5)        | 387.36    | 25.82 (15 joints)   |

training and testing sets are all real golf swings which are more complicated than the ones in [27] and without any limitation.

# 4 Conclusion and future work

A five chains full-body DBN model is proposed based on the spatial and temporal similarities of different golfers. A transportable, non-intrusive and inexpensive golf swing reconstruction system (SMRG) is built based on the DBN model. The experiments have shown that the system can reconstruct the golf swing with good quality. Although there are slight differences, compared with the MAT-T system, the reconstruction accuracy will increase with the development of depth imaging devices.

Our future work will be to incorporate more key joints (wrist and spine) or the club head movement into the system and analysis of kinematic parameters generated from the system. Reconstruction and analysis of other regular motions using depth imaging devices with our DBN model is also under consideration.

# References

1. Arvind D, Bates A (2008) The speckled golfer. In: Proceedings of the ICST 3rd international conference on Body area networks, Tempe, Arizona, USA. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), p 1–7
2. Chun S, Kang D, Choi H-R, Park A, Lee K-K, Kim J (2013) A sensor-aided self coaching model for uncocking improvement in golf swing. Multimedia Tools and Applications:1–27
3. Chun S, Kang D, Choi H-R, Park A, Lee K-K, Kim J (2014) A sensor-aided self coaching model for uncocking improvement in golf swing. Multimed Tools Appl 72(1):253–279
4. Coleman S, Anderson D (2007) An examination of the planar nature of golf club motion in the swings of experienced players. J Sports Sci 25(7):739–748
5. Evans K, Horan SA, Neal RJ, Barrett RS, Mills PM (2012) Repeatability of three-dimensional thorax and pelvis kinematics in the golf swing measured using a field-based motion capture system. Sports Biomech 11(2):262–272
6. Hadjiminas N, Child C (2012) Be The controller: a kinect tool kit for video game control. Paper presented at the The 5th Annual International Conference On Computer Games, Multimedia And Allied Technology (CGAT 2012), Bali, Indonesia
7. Holte MB, Chakraborty B, Gonzalez J, Moeslund TB (2012) A local 3-D motion descriptor for multi-view human action recognition from 4-D spatio-temporal interest points. IEEE J Sel Top Sign Proces 6(5):553–565
8. Karliga I, Hwang JN (2006) Analyzing human body 3-d motion of golf swing from single-camera video sequences. In: Acoustics, speech and signal processing, 2006. ICASSP 2006 Proceedings. 2006 I.E. International Conference on, Toulouse. IEEE, p V-V
9. Kenny I, Madden D, Downey J, Murray P, Campbell J, Breen S (2012) Biomechanical characterisation of leg movement during the golf swing following knee surgery. Paper presented at the 2012 British Association of Sport and Exercise Sciences Biomechanics Interest Group Conference
10. Kohli P, Shotton J (2013) Key developments in human pose estimation for kinect. Advances in Computer Vision & Pattern Recognition

11. Kwon YH (2009) Kinematic analysis of sports movements: golf swing plane analysis. In: ISBS-Conference Proceedings Archive, Limerick. p 31–34
12. Lin Y-H, Huang S-Y, Hsiao K-F, Kuo K-P, Wan L-T (2013) A kinect-based system for golf beginners' training. In: Information Technology Convergence, vol 253. Springer, p 121–129
13. Liu H, Wei X, Chai J, Ha I, Rhee T (2011) Realtime human motion control with a small number of inertial sensors. In: Symposium on interactive 3D graphics and games. ACM, p 133–140
14. Livingston MA, Sebastian J, Ai Z, Decker JW (2012) Performance measurements for the Microsoft Kinect skeleton. In: Virtual Reality Short Papers and Posters (VRW), 2012 I.E. Costa Mesa, CA. IEEE, p 119–120
15. MacKenzie SJ (2012) Club position relative to the golfer's swing plane meaningfully affects swing dynamics. Sports Biomech 11(2):149–164
16. MacKenzie SJ, Sprigings EJ (2009) A three-dimensional forward dynamics model of the golf swing. Sports Eng 11(4):165–175
17. Mann R, Griffin F, Yocom G (1998) Swing like a pro: the breakthrough scientific method of perfecting your golf swing. Random House Digital, Inc
18. MAT-T Website. http://tmpl.taylormadegolf.com/mat-t-system
19. McGuan SP (2002) Achieving commercial success with biomechanics simulation. In: 20 International Symposium on Biomechanics in Sports, Cáceres, Spain. p 451–460
20. Nesbit SM (2005) A three dimensional kinematic and kinetic study of the golf swing. J Sports Sci Med 4(4): 499–519
21. Nesbit SM, McGinnis R (2009) Kinematic analyses of the golf swing hub path and its role in golfer/club kinetic transfers. J Sports Sci Med 8(2):235–246
22. Obdržálek Š, Kurillo G, Ofli F, Bajcsy R, Seto E, Jimison H, Pavel M (2012) Accuracy and robustness of kinect pose estimation in the context of coaching of elderly population. Paper presented at the Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE San Diego, CA
23. Organization O Introducing OpenNI
24. Patsadu O, Nukoolkit C, Watanapa B (2012) Human gesture recognition using kinect camera. Paper presented at the Computer Science and Software Engineering (JCSSE), 2012 International Joint Conference on Bangkok
25. Seaman A, McPhee J (2012) Comparison of optical and inertial tracking of full golf swings. Procedia Eng 34:461–466
26. Shen W, Deng K, Bai X, Leyvand T, Guo B, Tu Z (2012) Exemplar-based human action pose correction and tagging. In: Computer Vision and Pattern Recognition (CVPR), 2012 I.E. Conference on, Providence, USA. IEEE, p 1784–1791
27. Shum HP, Ho ES, Jiang Y, Takagi S (2013) Real-time posture reconstruction for microsoft kinect. IEEE Trans Cybern 43(5):1357–1369
28. Smisek J, Jancosek M, Pajdla T (2011) 3D with Kinect. In: Computer Vision Workshops (ICCV Workshops), 2011 I.E. International Conference on Barcelona. IEEE, p 1154–1160
29. Suzuki S, Haake S, Heller B (2005) Skill analysis of the wrist release in golf swing to utilize shaft elasticity. In: Proceedings of APCST2005 - Asia-Pacific congress on sports technology. p 188–193
30. Urtasun R, Fleet DJ, Fua P (2005) Monocular 3D tracking of the golf swing. In: Computer vision and pattern recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. IEEE, p 932–938
31. Watanabe K, Hokari M (2006) Kinematical analysis and measurement of sports form. IEEE Trans Syst Man Cybern Part A Syst Humans 36(3):549–557
32. Wu Q, Boulanger P (2011) Real-time estimation of missing markers for reconstruction of human motion. In: Virtual Reality (SVR), 2011 XIII Symposium on, Uberlandia. IEEE, p 161–168
33. Zhang L, Hsieh J-C, Ting T-T, Huang Y-C, Ho Y-C, Ku L-K A (2012) Kinect based Golf Swing Score and Grade System using GMM and SVM. In: Image and Signal Processing (CISP), 2012 5th International Congress on, Chongqing. IEEE, pp 711–715
34. Zhang L, Hsieh JC, Wang J A (2012) Kinect-based golf swing classification system using HMM and Neuro-Fuzzy. In: Computer Science and Information Processing (CSIP), 2012 International Conference on, Xi'an. IEEE, p 1163–1166
35. Zhou H, Hu H (2008) Human motion tracking for rehabilitation—a survey. Biomed Signal Process Control 3(1):1–18

**Dongyue Lv** received the B.S. and M.S. in College of automation, Nanjing University of Aeronautics & Astronautics, China, in 2008 and 2011 respectively. He is currently working toward his Ph.D. degree from Institution of Electronics, Chinese Academy of Sciences. His research interests include computer vision, depth imaging and modeling.



**Zhipei Huang** received her Ph.D. degree from University of Chinese Academy of Sciences, China. She is an associate professor in the School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences. Her current research interests include signal detection and estimation, target tracking and computer vision.

**Lixin Sun** received the B.S. in College of Software, Nankai University, China in 2012. She is currently working toward her Ph.D. degree from School of Electronic, Electrical and Communication Engineering, Chinese Academy of Sciences. Her research interests include motion reconstruction, human neuromusculoskeletal modeling.



**Nenghai Yu** is currently a Professor in the Department of Electronic Engineering and Information Science of University of Science and Technology of China (USTC). He is the Executive Director of MOE-Microsoft Key Laboratory of Multimedia Computing and Communication, and the Director of Information Processing Center at USTC. He graduated from Tsinghua University, Beijing, China, and obtained his M.Sc. Degree in electronic engineering in 1992, and then he joined in USTC and worked there until now. He received his Ph.D. Degree in Information and Communications Engineering from USTC, Hefei, China, in 2004.

His research interests are in the field of multimedia information retrieval, digital media analysis and representation, media authentication, video surveillance and communications etc. He has been responsible for many national research projects. Based on his contribution, Professor Yu and his research group won the Excellent Person Award and Excellent Collectivity Award simultaneously from the National Hi-tech Development Project of China in 2004. He has contributed more than 80 papers to journals and international conferences.
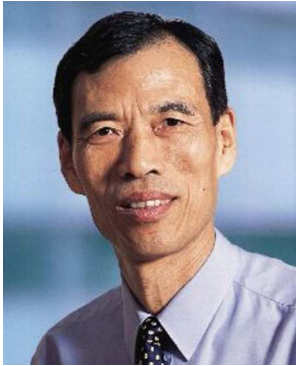
**Jiankang Wu** received the B.Sc. degree from the University of Science and Technology of China (UTSC) and the Ph.D. degree from Tokyo University, Japan. He is currently a professor in the University of Chinese Academy of Sciences, China. He is directing the Sensor Networks and Application Research Center, University of Chinese Academy of Sciences. Prior to the current position, he was the Principal Scientist and Department Manager of New Initiatives Department at Institute for Infocomm Research (I2R), Singapore, known as Kent Ridge Digital Labs (KRDL), from 1998 to 2001, and Institute of System Science (ISS), before 1998. He created new research initiatives in the areas of NeuroInformatics, PhysioInformatics, and embedded sensor network systems in I2R, and led a series of large international collaborative projects in KRDL and ISS. He was a Full Professor at the USTC until 1993.

Prof. Wu pioneered several researches in the area of visual information processing, including adaptive image coding in later 1970s, object-oriented GIS in early 1980s, face recognition technology in 1992, content-based multimedia indexing and retrieval in 1990s, and rights management for multimedia contents in later 1990s. He initiated and led three large international collaborative labs. He was regarded as "an exceptionally bright visionary" by the US Giga Information Group. He received nine distinguished awards from the nation and Chinese Academy of Sciences.