

文章编号:2095-6134(2016)02-0204-09

DBN model-based golf swing reconstruction using Kinect data^{*}

LÜ Dongyue¹, HUANG Zhipei^{2†}, TAO Guanhong², YU Nenghai¹, WU Jiankang²

(1 Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China;

2 University of Chinese Academy of Sciences, Beijing 100049, China)

(Received 20 November 2014; Revised 12 June 2015)

LY D Y, HUANG Z P, TAO G H, et al. DBN model-based golf swing reconstruction using Kinect data[J]. Journal of University of Chinese Academy of Sciences, 2016,33(2):204-212.

Abstract The last decade has witnessed an explosion of interest in golf, and the number of golf players has increased significantly. Therefore, how to train a golfer to make a proper and accurate golf swing has attracted extensive research attentions. Among these researches the most important step is swing capture and reconstruction. Thus far, restricted to the development of present depth imaging devices, of which the most famous one is Kinect, the initial captured swing movement may not be acceptable enough due to occlusions and mixing up of body parts. To restore motion information from self-occlusion and reconstruct 3D golf swing from low resolution depth data, a dynamic Bayesian network (DBN) model based golf swing reconstruction algorithm is proposed to increase the capture accuracy which integrates the spatial relationship among joints and their movement dynamics. Experimental results have proved that the proposed algorithm 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.

Key words golf swing reconstruction; dynamic Bayesian network model; Kinect

CLC number: TP391 **Document code:** A **doi:**10.7523/j.issn.2095-6134.2016.02.009

基于 DBN 模型与 Kinect 数据的高尔夫挥杆重建

吕东岳¹, 黄志蓓², 陶冠宏², 俞能海¹, 吴健康²

(1 中国科学院电子学研究所, 北京 100190; 2 中国科学院大学, 北京 100049)

摘要 人们对高尔夫运动的兴趣在过去的 10 年间爆发式增长, 同时高尔夫球员的数量也显著增加。因此, 如何训练一个球员做出适当和准确的挥杆动作已经引起研究者的广泛关注。在此类研究中最重要的一步是挥杆动作的捕获与重建。至今为止, 以 Kinect 为代表的深度成像设备受其基础条件的限制, 其捕获的挥杆运动可能会由于运动中肢体之间的互相遮挡与肢体识别时产生的混乱丧失足够的精确度。为了从肢体互相遮挡并且分辨率较低的深度图像信息中恢复比较精确的运动信息, 该文提出一个用于

* Supported by National Natural Science Foundation of China (61431017)

† Corresponding author, E-mail: communications2012@163.com

描述人体关节点之间空间关系与关节点动态特性的动态贝叶斯网络 (DBN) 模型, 并基于该模型实现了高尔夫挥杆动作的重建. 实验结果表明, 该算法能够实现重建精度媲美商用的光学运动捕获系统 (OMocap), 并且比现有的深度信息修改算法具有更好的性能.

关键词 高尔夫挥杆重建; 动态贝叶斯网络模型; Kinect

Golf is one of the most developed and popular sports. The number of active golf players is estimated at 50 million and golf course development has become the fastest growing property sector in the world. However, how to train a golfer to make a proper and accurate golf swing is critical to the success of the golf training course^[1]. Previously, the training was usually guided by professional golfers, which involved extensive human recourse. In recent years, there was a trend to use standard swing procedures for golf training, which considered many important parameters of swing, such as club head trace^[2-5], swing plane^[6], hub path^[7], leg movement^[8], and even wrist angle^[9-10]. However, among these parameters, the key is to accurately capture the swing movement. For this purpose, two kinds of motion capture (Mocap) systems are mainly used in recent years^[11]: optical Mocap (OMocap) system^[12-15] and wearable micro-sensors base Mocap system (MMocap)^[16-20]. The OMocap systems require to attach reflective markers on golfer's body segment, and then positions of the marker are obtained via multiple high-speed cameras around the golfer. The MMocap systems also require to place micro-sensors on golfer's body segments for motion reconstruction. Although these two kinds of Mocap systems have been used successfully in many applications, both systems require the golfer to wear some extra devices, which are intrusive and make the golfer uncomfortable during the swing movement.

Recently, newly introduced Microsoft Kinect, which is a contactless motion capture device using an RGB-D sensor providing synchronized color and depth images, has shown some potential to overcome the pitfalls of the traditional Mocap systems^[21]. The main challenge for the Kinect is its poor performance due to occlusions and mixing up of body parts^[22]. To tackle this problem and make it applicable for the

golf swing analysis, some ad-hoc solutions to improve the estimation accuracy have been proposed so far. For example, Zhang et al.^[23-24] used the Kinect to capture the 3D skeleton coordination of a golfer who was performing swing, and then used serial correlation model to score and recognize grade of golf swing. Lin et al.^[25] utilized the Kinect to capture the swing motion, and automatically identified the 6 common mistakes on swing motion. Shen et al.^[26] tried to tackle the occlusion problem and presented an example-based method to learn and correct the initially estimated poses from Kinect. Although these methods made some progress to improve the wrongly recognized swing posture from Kinect, the challenge is still not very well solved yet due to the ignorance of motion similarity contained in the swing dynamics.

However, there are similarities among swing movements performed by the same person: 1) spatial similarity (the relative movements among joints during swing movement are similar if the swings are performed without external interference) and 2) temporal similarity (the swing periods are similar). These similarities, if applied properly, can 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 this paper, due to the advantage of fusing spatial and temporal information, a dynamic Bayesian network (DBN) model-based golf swing reconstruction algorithm for Kinect is proposed. The algorithm integrates the spatial relationship among joints and their movement dynamics. The DBN model is trained using actual golf swing movement data. The joint positions generated from Kinect are optimized by the model. The reconstruction of golf swing is performed using these optimized positions. We have compared our

position outputs with the commercial OMocap system MAT-T^[27] and the method proposed in Ref. [26]. The good performance results have shown that the proposed algorithm achieves comparable tracking accuracy to the MAT-T system and improves the Kinect outputs much more than the methods mentioned in Refs. [22] and [26].

1 DBN model-based swing movement reconstruction

Generally, the human body can be regarded as a skeleton model shown in Fig. 1 (a). The skeleton model contains five chains. Each chain is a set of segments connected by joints. The joint positions are key elements to the whole swing procedure reconstruction. 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.

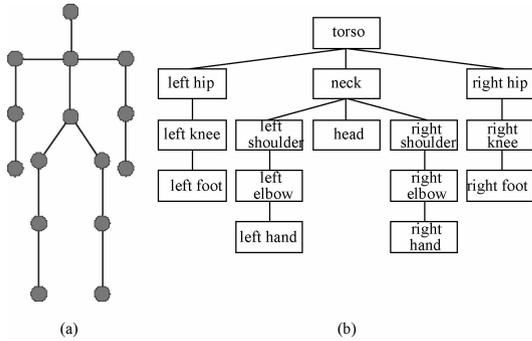


Fig. 1 Human skeleton model (a) and hierarchical structure (b)

The above golfer model includes 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. Their position information during time represents the whole body movement precisely. These 15 key joints construct a hierarchical structure of a golfer containing the five serial chains, which is shown in Fig. 1 (b). 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 in one arbitrary chain. It can be derived similarly in other four chains by using the

DBN model of the chosen chain.

1.1 Model representation

In the DBN model, 5 symbols are used to represent the states and observations:

s_t^i : The relative position of the i_{th} joint in chain in its parent joint's coordinate system (JCS) at time t .

X_t^i : The absolute position of the i_{th} joint in chain in global coordinate system (GCS) at time t .

Y_t^i : The position observation of the i_{th} joint in chain in GCS at time t .

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 is shown in Fig. 2. 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.

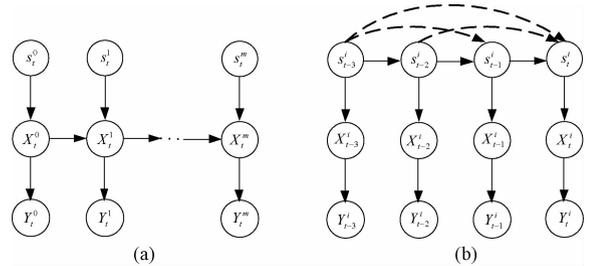


Fig. 2 DBN model of one arbitrary chain of human model: static structure (a) and dynamic structure (b)

The purpose of the DBN model is to estimate more precise positions of the joints in GCS during the whole swing, i. e., to 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

$$X_1^0, \dots, X_1^m, X_2^0, \dots, X_2^m, \dots, X_t^0, \dots, X_t^m. \quad (1)$$

This means reconstructing the swing of the golfer accurately, and 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-1}^{i-1} in the the same frame and its relative position s_{t-1}^i in the previous frame.

1.2 Model inference

As discussion above, the key to acquire ment of precise position of X_t^i is to calculate the posterior probability $\Pr(X_t^i | Y_{1:t}^{0:i})$, which is related to $\Pr(X_{t-1}^{i-1} | Y_{1:t}^{0:i-1})$ and $\Pr(s_{t-1}^i | Y_{1:t-1}^{0:i})$. This could be derived step by step from the joint probability distribution which can be written as

$$\Pr(s_{1:t}^{0:i}, X_{1:t}^{0:i}, Y_{1:t}^{0:i}) = \frac{\Pr(s_{1:t}^{0:i-1}, X_{1:t}^{0:i-1}, Y_{1:t}^{0:i-1}) \Pr(s_{1:t-1}^{0:i}, X_{1:t-1}^{0:i}, Y_{1:t-1}^{0:i})}{\Pr(s_{1:t-1}^{0:i-1}, X_{1:t-1}^{0:i-1}, Y_{1:t-1}^{0:i-1})} \cdot \Pr(X_t^i | X_{t-1}^{i-1}, s_t^i) \Pr(s_t^i | s_{t-1:t-n}^i) \Pr(Y_t^i | X_t^i). \quad (2)$$

Then, by doing parameter elimination, the probability contains only s_t^i and X_t^i and can be derived

$$\Pr(s_t^i, X_t^i, Y_{1:t}^{0:i}) = \frac{\Pr(Y_t^i | X_t^i)}{\Pr(Y_{1:t-1}^{0:i-1})} \left[\sum_{X_{t-1}^{i-1}} \Pr(X_{t-1}^{i-1}, Y_{1:t-1}^{0:i-1}) \Pr(X_t^i | X_{t-1}^{i-1}, s_t^i) \right] \cdot \left[\sum_{s_{t-1:t-n}^i} \Pr(s_{t-1}^i, Y_{1:t-1}^{0:i-1}) \Pr(s_t^i | s_{t-1:t-n}^i) \right]. \quad (3)$$

From (3), the posterior probability $\Pr(X_t^i | Y_{1:t}^{0:i})$ and $\Pr(s_t^i | Y_{1:t}^{0:i})$ can be derived

$$\Pr(X_t^i | Y_{1:t}^{0:i}) = C \cdot \Pr(Y_t^i | X_t^i) \cdot \sum_{X_{t-1}^{i-1}} \Pr(X_{t-1}^{i-1} | Y_{1:t-1}^{0:i-1}) \sum_{s_t^i} \Pr(X_t^i | X_{t-1}^{i-1}, s_t^i) \cdot \sum_{s_{t-1:t-n}^i} \Pr(s_{t-1}^i | Y_{1:t-1}^{0:i-1}) \Pr(s_t^i | s_{t-1:t-n}^i), \quad (4)$$

$$\Pr(s_t^i | Y_{1:t}^{0:i}) = C \cdot \sum_{s_{t-1:t-n}^i} \Pr(s_{t-1}^i | Y_{1:t-1}^{0:i-1}) \Pr(s_t^i | s_{t-1:t-n}^i) \cdot \sum_{X_t^i} \Pr(Y_t^i | X_t^i) \sum_{X_{t-1}^{i-1}} \Pr(X_{t-1}^{i-1} | Y_{1:t-1}^{0:i-1}) \Pr(X_t^i | X_{t-1}^{i-1}, s_t^i). \quad (5)$$

In (4) and (5), $C = \frac{\Pr(Y_{1:t}^{0:i-1}) \Pr(Y_{1:t-1}^{0:i})}{\Pr(Y_{1:t-1}^{0:i-1}) \Pr(Y_{1:t}^{0:i})}$ can be considered as a constant. Recalling the preliminary iteration sequence to estimate X_t^i in (1) and considering (4) and (5), a modified iteration sequence is proposed to estimate every joint position in chain at time t :

$$s_{t-1}^0, X_t^0, \dots, s_{t-1}^m, X_t^m. \quad (6)$$

1.3 Conditional probability distribution (CPD) definition

The calculations of $\Pr(X_t^i | Y_{1:t}^{0:i})$ and $\Pr(s_t^i | Y_{1:t}^{0:i})$ need three key elements: spatial relationship $\Pr(X_t^i | X_{t-1}^{i-1}, s_t^i)$, 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 the 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.

1.3.1 Spatial relationship

According to the human skeleton model, the joint positions are the only elements to be 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 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-1}^{i-1} , according to the chain structure. The spatial relationship between parent and child joints can be written as

$$\Pr(X_t^i | X_{t-1}^{i-1}, s_t^i) = N(X_{t-1}^{i-1} + s_t^i, Q_i). \quad (7)$$

In (7), Q_i is the process noise from previous training sets and their ground truths of one golfer.

1.3.2 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

$$\Pr(Y_t^i | X_t^i) = N(X_t^i, R_i). \quad (8)$$

1.3.3 Temporal relationship

For each golfer, the motion of joints should be unique and 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

$$\Pr(s_t^i | s_{t-1:t-n}^i) = N(A_{t-1}^i [s_{t-1}^i, \dots, s_{t-n}^i]^T, \Sigma_{t-1}^i). \quad (9)$$

In (9), the matrices A_{t-1}^i and Σ_{t-1}^i contain the parameters which need to be trained to get from $s_{t-1:t-n}^i$ to s_t^i .

2 Experiments and discussion

2.1 Setup

The MAT-T system and the algorithms in Refs. [22, 26] are used to evaluate the reconstruction performance of the proposed algorithm. The placement of the 6 cameras and the application environment are shown in Fig. 3, while the Kinect is put in front of the golfer at a distance of about 2.5 m with a height of about 1 m from the ground to make sure the whole swing motion can be captured. The joint position observations are provided by OpenNI skeleton functions^[28]. These functions can generate preliminary joint positions from the RGB-D information.



Fig. 3 The application environment of the MAT-T system

Five 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 of one golfer, all the 6 swings are used as testing set by turn. While 1 swing is testing set, the other 5 swings are training sets to generate “prior knowledge” of the DBN model. This ensures that every swing can have a corresponding reconstruction result. The two kinds of devices (i. e. 6 cameras MAT-T system and Kinect) capture the swing synchronously. The sampling rate of the cameras from

MAT-T system is 180 Hz, while the Kinect is 30 Hz.

To illustrate the feasibility of our algorithm, the joints and body segments which have the most significant movements or rotations were chosen. In our implementation, we chose the hand arms to evaluate the performance of the proposed algorithm. The shoulder width (SW) and arm 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 body parts can be acquired easily. By comparing with outputs of MAT-T system, we calculate the error ratio e between our outputs and MAT-T system outputs:

$$e = \frac{|L - \tilde{L}|}{L}, \quad (10)$$

In (10), L is the body part length from MAT-T system, while \tilde{L} is the body part length from our model.

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

$$\text{msJE} = \frac{\sum_{t=1}^T \sum_{i=1}^N |X_t^i - \tilde{X}_t^i|}{T}, \quad (11)$$

$$\text{mmsJE} = \frac{\sum_{t=1}^T \sum_{i=1}^N |X_t^i - \tilde{X}_t^i|}{T \times N}. \quad (12)$$

In (11) and (12), X_t^i is the position of joint i at time t from MAT-T, and \tilde{X}_t^i is from our model. N is the number of joints used to represent golfer (in our experiment $N = 15$). T represents the swing period, which equals the frame number Kinect captured during the whole swing.

2.2 Order number selection

Since the proposed model is multi-order DBN, a proper order number should be determined before the experiment. To evaluate the model performance, i. e. , to quantify the average accuracy on the whole

5 testing data, the msJE is calculated and shown in Fig. 4. The order number rises from 1 to 7, which causes the msJE to drop first and rise when the order number is above 5. For economizing training and calculating consumption, other order numbers are not considered. Apparently, the order number of our model is chosen to be 5.

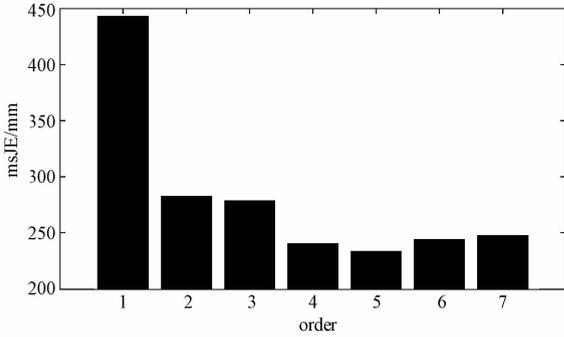


Fig. 4 Variation of msJE when the order of model rises

2.3 Computation consumption

The proposed DBN model is constructed and trained using the BNT (Bayesian network toolbox) in Matlab (R2011b). The joint positions are

modified offline with a speed of 10 fps (CPU: i5 - 2400 3.1 GHz).

2.4 Comparison with MAT-T

The hand position outputs of our algorithm, the MAT-T system, and the original Kinect are shown in Fig. 5. Because the joint moves in 3D space, the three components (x , y , and depth) of positions are shown.

In Fig. 5, the outputs of the former two kinds of results are only slightly different. 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 between the former 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 occur. These outputs will definitely affect the performance of the model.

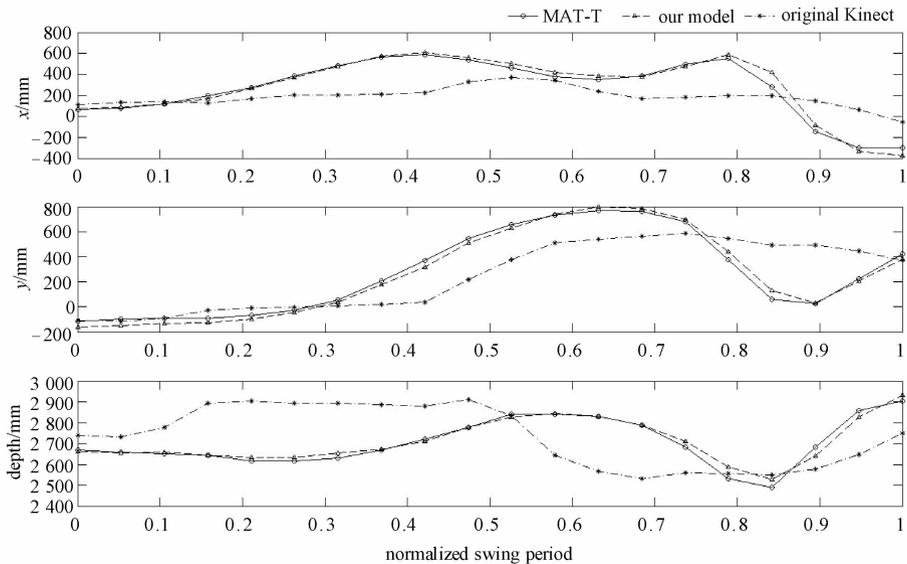


Fig. 5 Comparison of hand position outputs

The mentioned 5 body parts are chosen in the comparison. The results of comparison are shown in Table 1. The 5 body parts all have mean values of the whole swing procedure.

Compared with the MAT-T system, our model

do have certain errors in the outputs. Moreover, some body part's errors are significant (larger than 10%). These errors do 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 most rapidly. Markers may change their positions relative to their corresponding body segments due to skin movement^[29]. If the shoulder markers change positions, the outputs of the MAT-T system will fluctuate and this may cause the difference. The shoulder widths (SWs) of five golfers are shown in Fig. 6. The results are gained by our model, the MAT-T system, and original Kinect output.

In normal condition, the shoulder width of each golfer should keep constant during swing, because the body segments are rigid. In Fig. 6, the shoulder width results of the MAT-T system are not constant. The main reason is that the shoulder markers have changed their positions 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 algorithm show more robustness than the other two kinds of results. There are some differences between the outputs of our model 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), there is no doubt that the performance will

be better.

Table 1 Length comparison among body parts

| | body part | MAT-T/ | our model/ | e/% |
|----------|-----------|-----------|------------|-------|
| | | mm | mm | |
| Golfer 1 | SW | 235.458 4 | 265.052 1 | 12.57 |
| | LU | 250.218 6 | 228.117 3 | 8.83 |
| | LL | 234.175 7 | 247.271 2 | 5.59 |
| | RU | 250.218 9 | 233.607 2 | 6.64 |
| | RL | 234.175 2 | 252.271 2 | 7.73 |
| Golfer 2 | SW | 246.801 8 | 232.900 3 | 5.63 |
| | LU | 259.717 7 | 248.654 5 | 4.26 |
| | LL | 243.066 0 | 248.655 0 | 2.30 |
| | RU | 259.717 6 | 274.476 3 | 5.68 |
| | RL | 243.065 1 | 254.476 6 | 4.69 |
| Golfer 3 | SW | 269.044 9 | 305.361 4 | 13.50 |
| | LU | 303.933 4 | 279.851 3 | 7.92 |
| | LL | 284.447 4 | 272.659 4 | 4.14 |
| | RU | 303.933 6 | 273.988 4 | 9.85 |
| | RL | 284.445 7 | 302.122 3 | 6.21 |
| Golfer 4 | SW | 275.892 0 | 283.776 2 | 2.86 |
| | LU | 291.777 6 | 279.857 2 | 4.09 |
| | LL | 273.070 7 | 284.857 6 | 4.31 |
| | RU | 291.777 3 | 281.474 3 | 3.53 |
| | RL | 273.069 7 | 281.473 5 | 3.08 |
| Golfer 5 | SW | 274.379 1 | 281.775 9 | 2.70 |
| | LU | 297.526 7 | 289.945 1 | 2.54 |
| | LL | 278.450 2 | 259.123 0 | 6.94 |
| | RU | 297.526 2 | 291.855 7 | 1.91 |
| | RL | 278.449 1 | 275.661 1 | 1.00 |

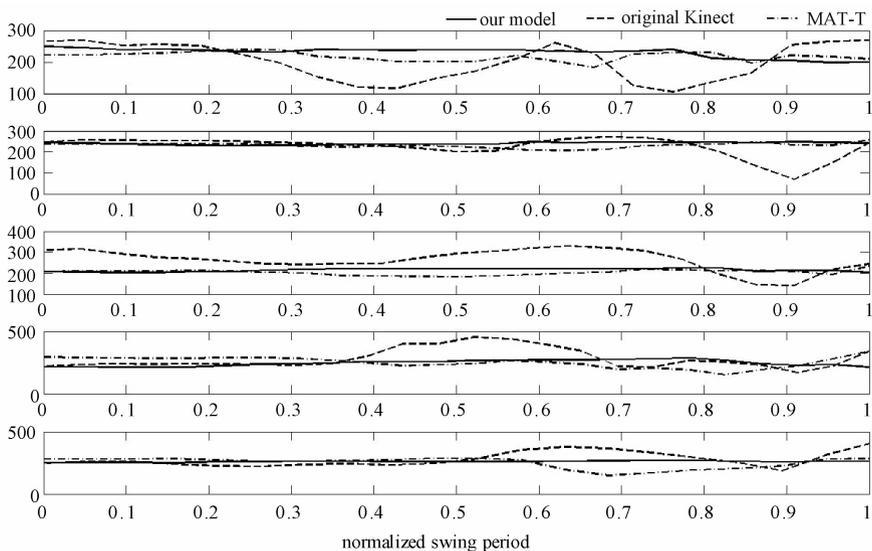


Fig. 6 The shoulder widths of five golfers during whole swing

2.5 Comparison with other algorithms

In order to show the superiority of our algorithm to the state of art Kinect based joint modification

algorithms, the algorithm in Ref. [22] and the RFR algorithm in Ref. [26] are chosen. The reasons why these algorithms are chosen as our opponent are two.

One is that though the rotation of golfer is limited, the test motions are golf swing and the other is that they provide comparable results with OMocap systems.

As mentioned above, the msJE was calculated. Because the numbers of the selected joints are different (15 in ours and 20 in the other two), a more precise value, the mean value of msJE (mmsJE), is also calculated. The mmsJE is the mean joint error of one joint during the whole swing. These two groups of values are shown in Table 2. The comparison shows that both the three algorithms can improve the Kinect performance but our algorithm performs much better.

Table 2 Comparison of msJE and mmsJE

| | msJE/mm | mmsJE/mm |
|---------------------|----------|---------------------|
| original | 2 071. 6 | 103. 58 (20 joints) |
| RFR algorithm | 1 575. 7 | 78. 79 (20 joints) |
| system in Ref. [22] | 1 452. 0 | 72. 60(20 joints) |
| our algorithm | 233. 65 | 15. 58 (15 joints) |

By comparing the outputs of our model, the MAT-T system, and two other algorithms, We find that our algorithm works well with high robustness against severe joint occlusion and improves the Kinect performance. The outputs of our model are comparable with the MAT-T system, which illustrates the feasibility and effectiveness of the proposed golf swing reconstruction algorithm. The comparison between our algorithm and two other algorithms shows that the mmsJE of our outputs is much lower. Moreover, our training and testing sets are all real golf swings without any limitation.

3 Conclusion and future work

A golf swing reconstruction algorithm is proposed based on Kinect and a full-body DBN model. The experiments have shown that it can reconstruct the golf swing with good quality. Although there are slight differences compared with the MAT-T system, the accuracy of the proposed algorithm will increase with the development of hardware and rise of sampling rate.

Our future work will be to incorporate the club

head into the model and analysis of kinematic parameters generated by the model. Reconstruction and analysis of other regular motion using Kinect with our algorithm is also under consideration.

References

- [1] Mann R, Griffin F, Yocom G. Swing like a pro: the breakthrough scientific method of perfecting your golf swing [M]. Random House Digital, Inc. ,1998 :
- [2] Urtasun R, Fleet D J, Fua P. Monocular 3D tracking of the golf swing [C] // Computer Vision and Pattern Recognition 2005. IEEE Computer Society Conference on. IEEE,2005 : 932-938.
- [3] Karliga I, Hwang J N. Analyzing human body 3-d motion of golf swing from single-camera video sequences [C] // Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on: Toulouse IEEE,2006:V-V.
- [4] MacKenzie S J. Club position relative to the golfer's swing plane meaningfully affects swing dynamics [J]. Sports Biomechanics, 2012, 11(2) :149-164.
- [5] Coleman S, Anderson D. An examination of the planar nature of golf club motion in the swings of experienced players[J]. Journal of Sports Sciences,2007,25(7) :739-748.
- [6] Kwon Y H. Kinematic analysis of sports movements: golf swing plane analysis [C] // ISBS-Conference Proceedings Archive: Limerick; 2009 : 31-34.
- [7] Nesbit S M, McGinnis R. Kinematic analyses of the golf swing hub path and its role in golfer/club kinetic transfers [J]. Journal of Sports Science and Medicine, 2009, 8(2) : 235-246.
- [8] Kenny I, Madden D, Downey J, et al. Biomechanical characterisation of leg movement during the golf swing following knee surgery[C] //2012 British Association of Sport and Exercise Sciences Biomechanics Interest Group Conference. 2012.
- [9] Chun S, Kang D, Choi H-R, et al. A sensor-aided self coaching model for uncocking improvement in golf swing[J]. Multimedia Tools and Applications, 2014, 72(1) : 253-279.
- [10] Suzuki S, Haake S, Heller B. Skill analysis of the wrist release in golf swing to utilize shaft elasticity [C] // Proceedings of APCST2005-Asia-Pacific congress on sports technology. 2005: 188-193.
- [11] Zhou H, Hu H. Human motion tracking for rehabilitation: A survey[J]. Biomedical Signal Processing and Control, 2008, 3(1) : 1-18.
- [12] Holte M B, Chakraborty B, Gonzalez J, et al. A local 3-D motion descriptor for multi-view human action recognition from 4-D spatio-temporal interest points [J]. Selected Topics in Signal Processing, IEEE Journal of, 2012, 6(5) : 553-565.

- [13] MacKenzie S J, Sprigings E J. A three-dimensional forward dynamics model of the golf swing [J]. *Sports Engineering*, 2009, 11(4): 165-175.
- [14] Wu Q, Boulanger P. Real-time estimation of missing markers for reconstruction of human motion [C] // *Virtual Reality (SVR)*, 2011 XIII Symposium on. *Uberlandia*: IEEE, 2011:161-168.
- [15] Nesbit S M. A three dimensional kinematic and kinetic study of the golf swing[J]. *Journal of Sports Science and Medicine*, 2005, 4(4): 499-519.
- [16] Watanabe K, Hokari M. Kinematical analysis and measurement of sports form [J]. *Systems, Man and Cybernetics, Part A: Systems and Humans*, IEEE Transactions on, 2006, 36(3): 549-557.
- [17] Arvind D, Bates A. The speckled golfer[C]//*Proceedings of the ICST 3rd international conference on Body area networks*: Tempe, Arizona, USA. 2008: 1-7.
- [18] Liu H, Wei X, Chai J, et al. Realtime human motion control with a small number of inertial sensors [C] // *Symposium on Interactive 3D Graphics and Games*: ACM. 2011: 133-140.
- [19] Seaman A, McPhee J. Comparison of optical and inertial tracking of full golf swings[J]. *Procedia Engineering*, 2012, 34: 461-466.
- [20] Evans K, Horan S A, Neal R J, et al. Repeatability of three-dimensional thorax and pelvis kinematics in the golf swing measured using a field-based motion capture system [J]. *Sports Biomechanics*, 2012, 11(2): 262-272.
- [21] Smisek J, Jancosek M, Pajdla T. 3D with Kinect [C] // *Computer Vision Workshops (ICCV Workshops)*, 2011 IEEE International Conference on Barcelona IEEE. 2011: 1 154-1 160.
- [22] Shum H P, Ho E S, Jiang Y, et al. Real-time posture reconstruction for Microsoft Kinect [J]. *Cybernetics, IEEE Transactions on*, 2013, 43(5): 1 357-1 369.
- [23] Zhang L, Hsieh J-C, Ting T-T, et al. A Kinect based golf swing score and grade system using GMM and SVM [C] // *Image and Signal Processing (CISP)*, 2012 5th International Congress on. *Chongqing*: IEEE, 2012: 711-715.
- [24] Zhang L, Hsieh J C, Wang J. A Kinect-based golf swing classification system using HMM and Neuro-Fuzzy [C] // *Computer Science and Information Processing (CSIP)*, 2012 International Conference on. *Xi'an*: IEEE, 2012: 1 163-1 166.
- [25] Lin Y-H, Huang S-Y, Hsiao K-F, et al. A Kinect-based system for golf beginners' training [M]. *Springer*, 2013: 121-129.
- [26] Shen W, Deng K, Bai X, et al. Exemplar-based human action pose correction and tagging [C] // *Computer Vision and Pattern Recognition (CVPR)*, 2012 IEEE Conference on. *Providence, USA*: IEEE, 2012: 1784-1791.
- [27] MAT-T Website [DB/OL]. [2014-11-23]. <http://tmpl.taylormadegolf.com/mat-t-system>.
- [28] Introducing OpenNI [M/OL]. (2013-12-12)[2014-10-30]. http://www.openni.org/images/stories/pdf/OpenNI_UserGuide.pdf.
- [29] McGuan S P. Achieving commercial success with biomechanics simulation [C] // *20 International Symposium on Biomechanics in Sports*. *Cáceres, Spain*: 2002: 451-460.