Augmented Tai-Chi Chuan Practice Tool with Pose Evaluation

[1]Yao-Fu Jan, [2]Kuan-Wei Tseng, [2]Peng-Yuan Kao, [1,2]Yi-Ping Hung
[1]Department of Computer Science and Information Engineering,
[2]Graduate Institute of Networking and Multimedia,
National Taiwan University, Taipei, Taiwan
hung@csie.ntu.edu.tw

***Abstract*—Tai Chi Chuan (TCC) is a well-known Chinese martial art that promotes health. In addition to learning TCC from a coach in a classroom setting, learners usually use books or videos to practice on their own. However, since turning is a frequent movement in TCC, learners cannot watch a tutorial and practice TCC at the same time. Furthermore, it is difficult for users to determine whether their postures are correct. We propose an augmented reality TCC practice tool with pose evaluation to help people practice TCC on their own. The tool consists of an optical see-through head-mounted display, external cameras, digital compasses, and a server. Users learn TCC movements from surrounding virtual coaches in augmented reality and determine whether their postures are correct via an evaluation module. Study results show that the proposed tool provides a helpful learning environment for TCC and that the pose estimation and evaluation are robust and reliable.**

***Index Terms*—Tai-Chi-Chuan, Sport Learning, Pose Estimation, Augmented Reality, Human Computer Interaction**

I. Introduction

Many studies show that Tai Chi Chuan (TCC) improves health both physically and mentally [1]–[5]. Usually, TCC is learned from a coach in a class setting. As coaches are not available outside of class time, people typically consult instruction books or watch demonstration videos to practice TCC on their own.

Nevertheless, there are several drawbacks to such books and videos. First, people generally see the coach's posture from a specific, fixed point of view; they cannot change to whatever point of view would be better. Second, as some TCC movements involve turning the head, checking one's posture in such a case obliges the student to stop performing the movement before consulting the instruction book or demonstration video, which interrupts their progress. Last, it is difficult for users to determine whether their postures and movements are correct. Although looking into a mirror is a solution, few users have the luxury of a dance classroom environment. Thus the perspective or occlusion problem still exists.

To help people practice TCC on their own, we propose a TCC practice system consisting of a learning module and an evaluation module. In the learning module, users wear an optical see-through head-mounted display (OST-HMD) and see virtual coaches surround them in augmented reality (AR). Fig. 1 shows the architecture of the proposed system. AR and the HMD allow the user to simultaneously observe coaches from multiple perspectives and practice the movement without distractions or interruptions. The evaluation module helps users to determine whether their movements or postures are correct. With digital compasses attached to the feet and external camera(s) continuously monitoring users' feet orientation and capturing imagery, we reconstruct the user skeleton at the server and compare it with the correct movements to evaluate their performance. We summarize the key functionalities of this work:

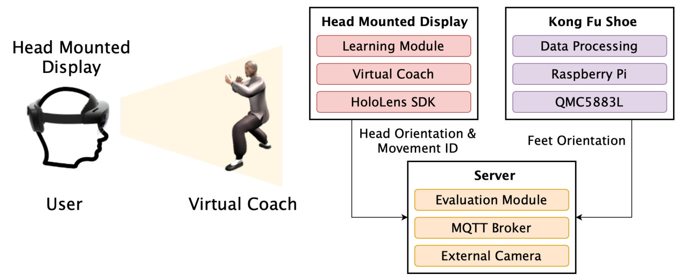


Fig. 1. Scenario drawing and the system architecture of the proposed augmented Tai-Chi Chuan practicing tool

- **Virtual Coach**: The user inspects and imitates coach movements from multiple perspectives to practice TCC in augmented reality. (Section III.A)
- **Movement Learning**: To aid users when practicing continuous movements, we label all key postures in every movement. (Section III.B)
- **Pose Estimation**: We analyze color images from the camera and digital compass measurements to reconstruct the 3D human pose with feet orientation. (Section IV.A)
- **Pose Evaluation**: To highlight differences between the user's pose and that of the coach, we align and compare the postures, output the corresponding scores, and offer video for review. (Section IV.B)

The contributions of this work are threefold. First and foremost, we propose a versatile augmented TCC learning module. We create virtual coach of Yang-Style TCC which surrounds the user in AR based on real TCC coach's movement. Secondly, we propose a pose evaluation module to help user improve their practice. We design a novel grading rubrics along with review interface to remind users of noteworthy actions. Last but not least, from our user experience study, we investigate user preference on the layout of virtual coaches and interaction method in AR system, which serve as reference for subsequent development of AR sport learning system.

II. Related Work

A. Sport Learning Systems

When learning or practicing sports at home, a flat display such as a television or projector is the most common multimedia. MotionMA [6] and YouMove [7] use Kinect to analyze and model experts' motions. Comparison to student movement yields feedback information that is shown on the screen. Physio@Home [8] uses a wearable device and two cameras that view the user from different angles. Via guidance information in front of a TV, the system decreases the number of mistakes when physiotherapy patients do their exercises through observation from different angles. SleeveAR [9] also shows the potential of projector-based guiding systems in guiding user hand movements. Recent years have seen increased use of OST-HMDs. Stylo-Handifact [10] consists of a haptic device attached to the forearm and a visualization of a virtual hand, which are combined to provide visuo-haptic feedback for posture training applications. Ikeda et al. [11] and Yeo et al. [12] use OST-HMDs for self-paced sport training to help the user imitate the expert's motions.

B. Tai Chi Chuan Learning Systems

Chen et al. [13] use pressure-sensing insoles for learning TCC. They use a smartphone or tablet to show visual feedback on the weight distribution of the user's feet as well as the virtual coach's feet. "Tai Chi in the Clouds" [14] uses a micro unmanned aerial vehicle (UAV) as "clouds" to lead or follow hand movements, providing live feedback on smoothness of movement via LEDs. Chua et al. [15] propose a training system for TCC based on a VR headset to guide body movements using virtual coaches. Users watch their body information in the virtual world via an avatar reconstructed from their skeleton animation. Whereas the system proposed by Iwaanaguchi et al. [16] is similar to that of [15], but their system only contains the virtual coach's movements.

In previous work [17], [18] we use an OST-HMD to learn TCC. Users easily evaluate their movements in a virtual mirror shown on the OST-HMD when practicing TCC. We also use wearable devices to detect wrist motion [19] and feet pressure [20], [21], which helps users to modify their body pose more effectively.

C. Pose Estimation and Evaluation

OpenPose [22] is a well-known real-time pose estimation method to detect multiple people's 2D poses from a single RGB image. It is a bottom-up method, which means that it detect joints first and then connects the joints to form the skeleton of the human body. Lifting from the Deep [23] is a deep architecture for detecting 2D and 3D pose from a single RGB image. It uses a four-stage architecture which optimizes 2D and 3D poses together to yield a refined 2D and 3D pose after a few iterations.

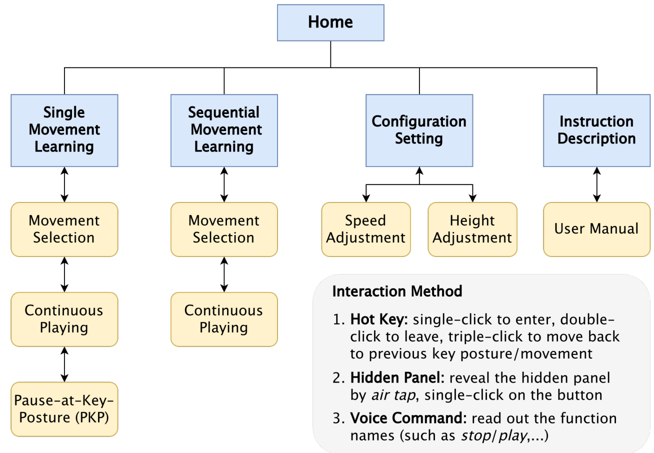


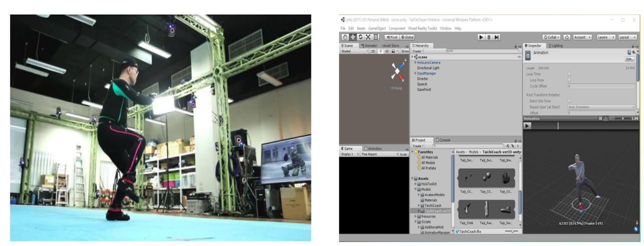
Fig. 2. Flow of user interaction for learning module

To evaluate pose correctness, we measure the pose similarity between the user and the coach. This has been measured by joint positions' Euclidean distance [7], [24], or in terms of the base of joint angles [6], [25]. Measuring similarity based on joint angles rather than joint positions removes the need for calibration procedures to eliminate the error caused by differences between users' body shapes.

III. Learning Module

In this section, we introduce the augmented reality TCC learning module of the proposed system. Fig. 2 shows the interaction flow of the learning module. The only hardware is the Microsoft HoloLens (first generation), a mixed-reality head-mounted display running the Windows Mixed Reality platform, which is equipped with multiple sensors, including inertial measurement units (IMUs), environment understanding cameras, a depth camera, and so on, which contribute to its robust environment localization and mapping.

A. Virtual Coach



(a) Movement recording (b) Animation creation

Fig. 3. Virtual coach creation

1) Virtual Coach Creation: There are many styles of TCC. In this work, we invited a TCC master to demonstrate standard Yang-Style TCC. His movements were recorded by a Vicon motion capture system consisting of 54 Vicon T160 cameras to form a capture volume of 15m (L) $\times$ 5m (W) $\times$ 3m (H). The movements recorded from the TCC master were first converted to the animation file and then imported into Unity 3D as the

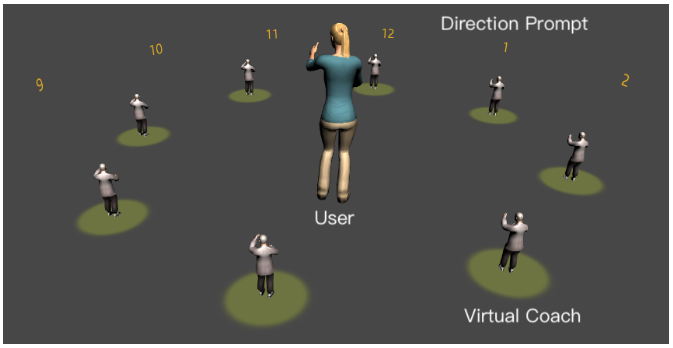


Fig. 4. Virtual coach layout in AR

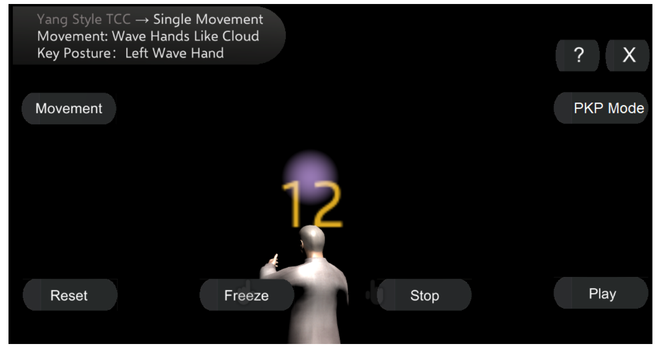


Fig. 5. Hidden panel in learning module

animation of the virtual coaches in our system. Fig. 3 shows the creation of the virtual coach.

2) Virtual Coach Layout: The user is surrounded by eight (this is customizable) virtual coaches in the augmented environment, as shown in Fig. 4. Users observe the motion not only of coaches in front of them but also of the coaches behind them. This solves the fixed perspective and occlusion problems. In addition, twelve numbers surround the user. These numbers are the direction prompts showing the clock position. Each movement starts from the twelve o'clock direction. The virtual coaches' orientations and the twelve numbers are set to the orientation of the user's head when the user enters the learning module, and can be changed by the user. While practicing, the relative distance between the virtual coaches and the user remains constant. However, since the horizontal and vertical field of view (FOV) of Microsoft HoloLens are restricted to 30 and 17.5 degrees respectively, we shrink the size of the virtual coaches to fit the FOV, which makes it slightly less realistic.

3) Configuration Setting: Here the user configures the environment. Users can adjust the demonstration speed and the mounting height of the virtual coach. The default (normal) speed is as fast as the TCC master. Users can adjust the speed according to their learning status. As for height adjustment, since user head shapes differ, the viewing angle varies slightly when wearing the HMD. Thus, the preset height of the virtual coach may not apply to all users. Users use the up and down buttons to adjust the height of the virtual coach positions.

B. Key Functionalities

There are two primary learning modes in the proposed learning module: single movement learning and sequential movement learning. Specifically in single movement learning, there are two playing modes: continous playing mode and pause-at-key-posture (PKP) playing mode. User can choose the appropriate mode according to their own learning status.

1) Single Movement Learning: There are 108 movements in Yang-Style TCC, each of which is composed of several key postures with distinctive names. In single movement learning, the user practices either key postures or a single movement in a continuous manner. After entering this mode, the user first selects one movement by name to practice from the movement selection menu. The name of the mode, movement, and key posture are shown in the top-left corner of the FOV in the HMD. Next, the user is surrounded by virtual coaches, whose spatial layout was specified in the previous section. All the coaches repeatedly demonstrate the movements that user has selected, while the system first speaks the name of the movement and then speaks every key posture as the coaches demonstrate it. The user follows the instructions and imitates the coaches' posture. If the user feels difficult in continous playing mode, he/she can enter PKP mode to practice a certain key posture. In this mode, coaches demonstrate the same movement but pause at each key posture for the user to learn and observe. For instance, there are three key postures in *Wave hand like cloud* (分勁雲手): *Unfolding palm* (撐手), *Left wave hand* (左雲手), and *Right wave hand* (右雲手). If the user practices this movement in PKP mode, the coaches pause at *Unfolding palm* before the user switches to the next key posture.

2) Sequential Movement Learning: Connection between each movement is vital in TCC, and thus experienced TCC practitioners prefer sequential movement. They are familiar with all TCC movements and tend to practice different movements continuously without interruptions. In sequential movement learning, the user either starts from the first TCC movement, i.e., *Preparation* (太極起式) or specifies any desired movement to start with. As with the single movement mode, the user is surrounded by virtual coaches who demonstrate the selected movements sequentially once. During the demonstration, the user can still switch to the next or previous movement.

C. User Interaction

To enhance the user experience, we provide three interaction methods: hidden panel, hot key, and voice command. In single or sequential movement learning, there are several buttons in the hidden panel, as shown in Fig. 5. These are invisible during practice but can be revealed by an *air tap*, which is a HoloLens gesture. The user can also control the system with the hot key, which can be triggered by a HoloLens gesture or the physical clicker. Generally speaking, the user single-clicks to move to the next movement or key posture, double-clicks to return to the parent node, and triple-clicks to move back to the previous movement or key posture. The user can also read out the command name to manipulate the system, for instance saying *stop* to pause the demonstration or *play* to resume.

Fig. 6. Evalution module layout. (We support one or multiple cameras.)

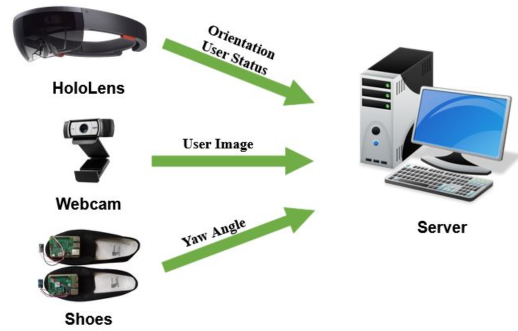


Fig. 7. Hardware and communication (transmit via MQTT protocal)

IV. Evaluation Module

In this section, we describe the evaluation module of the proposed system, both hardware and software. Fig. 6 shows the layout of the module: it consists of external cameras, a pair of special shoes equipped with a digital compass and single-board computers, and a remote server, as shown in Fig. 7. Image and measurement data are sent to the server, where we reconstruct the human pose and calculate the similarity to the coach's movement, which is taken as the ground truth for the scores and review video.

A. Digital Compasses

We installed a 3-Axis Magnetic Sensor QMC5883L and a Raspberry Pi single-board computer on a traditional kung-fu shoe. The Raspberry Pi receives the yaw angle measurements for both of the user's feet $\hat{\theta}_{LF}$ and $\hat{\theta}_{RF}$ at 5 Hz from the digital compasses and sends these data to the server wirelessly. To eliminate orientation bias between the learning system and the digital compass, we take the mean value of the yaw angle measurements as the base number and subtract each measurement from their mean to get the relative difference $\hat{\theta'}_{LF}$ and $\hat{\theta'}_{RF}$. We do the same with the ground-truth value to get θ'_{LF} and θ'_{RF}.

B. Pose Estimation with External Cameras

In pose estimation, we reconstruct the human pose by analyzing imagery from the external camera. If there are multiple cameras, we select one optimal perspective and discard the others. The image frames captured from the optimal perspective are fed as input to reconstruct the human pose. We estimate the 2D pose and use this to infer the 3D pose. For 2D pose estimation, we use OpenPose [22], a well-known bottom-up pose estimation method which detect joints, which are then connected to form the skeleton of the human body. First, it takes an input color image to estimate part confidence maps and part affinity fields (PAFs). A part confidence map describes the confidence of the locations of the human joint positions, whereas PAFs describe the relationship between each detected human joint. Then, OpenPose uses bipartite matching to connect the parts (joints) to form the skeleton. After acquiring the 2D pose, we follow Lifting from the Deep [23] for 3D pose estimation: we feed the 2D human pose into the pretrained probabilistic model to lift the 2D pose into a 3D pose. Then we extract the elbow angles $\hat{\theta'}_{LE}$, $\hat{\theta'}_{RE}$ and knee angles $\hat{\theta'}_{LK}$, $\hat{\theta'}_{RK}$ from the generated 3D pose.

C. Pose Evaluation

We align the measurement and ground-truth data via dynamic time warping (DTW) and measure the similarity between two sequences with the Gaussian function. The output is a score along with a review video, as shown in Fig. 10. We color correct, acceptable, and incorrect postures as green, yellow, and red so that the user can use the color to determine whether his/her posture is correct.

1) Dynamic Time Warping (DTW): Dynamic time warping is a fundamental algorithm for similarity measurements between two unequal-length signals. Let $\mathbf{G} = (g_1, g_2, \ldots, g_n)$ and $\mathbf{U} = (u_1, u_2, \ldots, u_m)$ denote the discrete time signal whose lengths are n and m respectively. The objective is to find a warping path $\mathbf{W} = \{(p_i, q_i) \mid 1 \le i \le k,\ k = \min(m, n)\}$ such that the warping cost D is optimized by dynamic programming:

$$D = \min_{\mathbf{W}} \sum_{i=1}^{k} \|\mathbf{G}(p_i) - \mathbf{U}(q_i)\|^2 \tag{1}$$

2) Similarity Measurement: Although the cumulative Euclidean distance calculated in DTW can represent the similarity between two sequences, it is not normalized. Hence, we apply a Gaussian function-based similarity metric to compare the aligned signal as

$$S_i = a \exp\left(-\frac{\|\mathbf{G}(p_i) - \mathbf{U}(q_i)\|^2}{2\sigma^2}\right), \tag{2}$$

where a is the normalization factor and σ is the Gaussian RMS width, which represents the tolerance of error.

3) Grading Rubric: Our grading rubric has three criteria: absolute elbow angles S_{elbow}, absolute knee angles S_{knee}, and relative foot angles S_{foot}. For each criteria, the score is calculated by Eq. 2. Note that we calculate the left side and right side separately. If the score of a certain criteria is less than 40, we color it red in the review video, implying the posture is incorrect. If the score lies between 40 and 60, we color it yellow, indicating acceptable posture. If the score is higher than 60, we color it green, showing the posture is correct. The

(a) Poor performance

(b) Good performance

Fig. 8. User interface of review system (green: correct; yellow: acceptable; red: incorrect)

total score, which represents the overall performance of the user, is the weighted sum of each criteria:

$$S_{\text{total}} = \sum_{i \in \{\text{elbow, knee, foot}\}} w_i S_i \tag{3}$$

where w_i and S_i denote the weight and score of each criteria.

V. Learning Module User Experience

A. Interaction Methodology

From the user's perspective, voice commands are convenient. Nevertheless, the built-in voice recognition system of the HoloLens (1st gen) is not very robust. It tends to fail due to environmental noise and user accents. Thus, the other two interaction methods—the hidden panel and the hot key—are indispensable. The hot key is triggered by physical clickers or HoloLens gestures. The user can single-click, double-click, or triple-click to manipulate the system. The hidden panel, in turn, contains several function buttons and can be revealed by an air tap. Each method has its advantages and disadvantages. In this part, we describe a user study conducted for comparison.

1) Experiment Procedure: Participants experimented with the two interaction methods with the HoloLens (1st gen). Before they started, we explained how the methods work. Then they followed our instructions to try out the functions in the learning module, after which they rated the operation methods and the functions using 7-point Likert scales.

2) Participants: Thirty six participants, including 23 females and 13 males, age 19 to 32, mean age 20.69 (*SD* = 2.19), were recruited for the study. Most participants come from Tai-Chi Chuan Beginning Class.

3) Result: Fig. 9(a) shows that the hidden panel was judged significantly better than the hot key in terms of *Continuous/PKP(Pause-at-Key Posture) Playing Mode Switch* (hotkey: mean = 4.75, *SD* = 1.11; hidden panel: mean = 5.00, *SD* = 1.39; one-tailed p = 0.40), *Manipulation in PKP Mode* (hotkey: mean = 4.00, *SD* = 1.60; hidden panel: mean = 5.17, *SD* = 1.50; one-tailed p = 2.158E-3), and *Manipulation in Sequential Movement Learning* (hotkey: mean = 3.89, *SD* = 1.51; hidden panel: mean = 5.17, *SD* = 1.42; one-tailed p = 4.31E-4). Participants commented that "triple-click is extremely difficult to access" (eighteen users) and "it takes times to memorize all hot key operation." (nine users).

TABLE I
Selected Yang-Style TCC Movements

Movement	Arm movement	Body rotation	Foot movement
Wave hand like cloud	X		
Grasp bird's tail	X	X	
Right brush knee, twist step	X	X	X

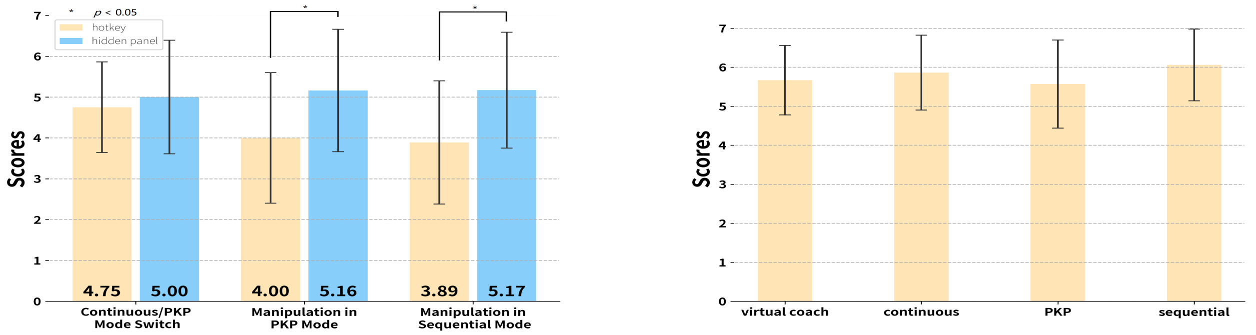
B. Overall Satisfaction

We also asked the participants about the overall user experience. "Usefulness of virtual coach (*virtual coach*)", and three modes "Continuous playing mode in Single Movement Learning (*continuous*)", "PKP mode in Single Movement Learning (*PKP*)", and "Sequential Movement Learning (*sequential*)" are evaluated. Most participants thought highly of the proposed learning module, as shown in Fig. 9(b). Although most were satisfied, some still suggested improvements: "it would be better to allow users to modify the orientation of coach" (eight users), "the buttons on hidden panel should be more centralized." (six users).

VI. Evaluation Module User Experience

A. Utility of Evaluation Module

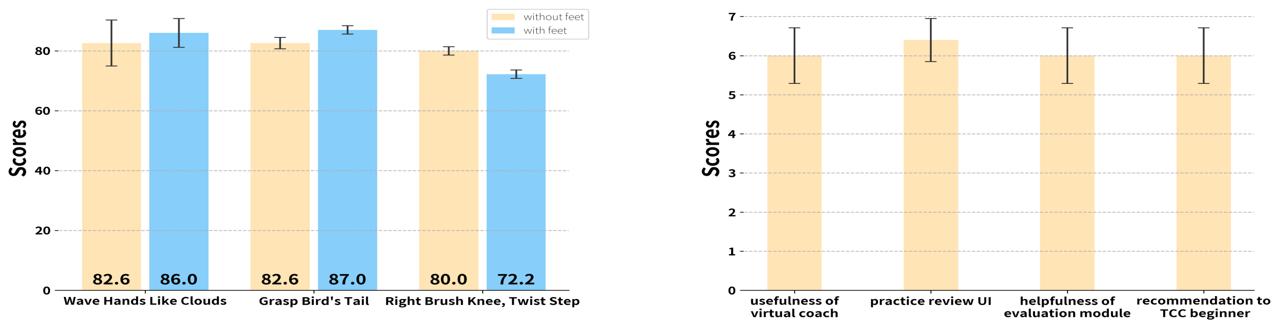
To verify the effectiveness of the proposed evaluation module, we invited five advanced TCC students to participate in user studies. Since it is unreasonable to ask TCC beginners or people who do not practice TCC to evaluate the pose estimation results, we could only depend on a TCC master or an experienced TCC practitioner. We selected three representative movements in Yang-Style TCC: *Wave hand like cloud* (分勁雲手), *Grasp bird's tail* (攬雀尾), and *Right brush knee, twist step* (右摟膝拗步). The difference between these movements is summarized in Table 1. These three movements employ different motion styles and represent different levels of difficulty, so we can evaluate the module's pose estimation and feet orientation estimation separately.



(a) Comparison between hot key and hidden panel for learning module

(b) User preference for each components in learning module

Fig. 9. User experience on learning module (36 participants; all of them are TCC beginners)



(a) Comparison on grading rubrics (w/wo feet orientation)

(b) Evaluation module user preference

Fig. 10. User experience on evaluation module (5 participants; all of them are experienced TCC practitioners)

1) Experiment Procedure: Before the experiment started, to familiarize the participants with the equipment, they were asked to wear the special shoes and HoloLens to practice TCC several times. Afterwards, they were asked to practice the selected TCC movements. During this practice, they were to asked to perform the movement incorrectly to check whether the evaluation module highlighted the erroneous pose.

2) Participants: Five male participants aged 24 to 26 with a mean age of 24.6 (SD = 0.89) were recruited for the study. All participants had experience learning TCC: one for 6 years, three for 5 years, and one for 4 years.

3) Result: We compared the scores using only the camera versus combining the camera and the digital compasses. Fig. 10(a) shows that for movements without any feet movement, the scores are higher with feet orientation. This results from the total score is the weighted sum of three criteria. Users can get nearly full scores of the relative foot angles S_{foot} by standing in place. However, for difficult movements with feet movements, the scores dropped. This seems to indicates that feet orientation is easily ignored when practicing strenuous movements. Hence, our grading rubric reminds the user to focus more on feet orientation.

B. User Preference

Next, participants were asked to use a 7-point Likert scale to rate the operability of the practice review user interface and to judge whether the evaluation module helped them to modify their movements. Also, participants were asked whether they would recommend TCC beginners to practice TCC on their own with our system. The user preferences are displayed in Fig. 10(b). Experienced TCC students agreed that the Evaluation Module clearly represented coach movements, that the practice review UI was easy to use, and that the module correctly identified errors and correct poses. Also, they believed that the proposed module would satisfy the needs of TCC beginners when memorizing TCC movements. One participant, however, commented that “This module did not clearly explain the reasons for errors and did not give directions on how to improve” (P4).

VII. Conclusion

We present an augmented Tai Chi Chuan practice tool with the following contributions: a learning module that provides labeled key postures and a user-friendly interaction method; an evaluation module which helps users to review and correct their movements; an probe into AR interaction methodology from user study. Our system architecture can also be applied to other sports that require close attention to posture. Despite that most users were satisfied with the system, we will continue to explore how to effecrtively guide the user to correct posture.

References

[1] C. Lan, J.-S. Lai, and S.-Y. Chen, "Tai chi chuan," *Sports medicine*, vol. 32, no. 4, pp. 217–224, 2002.

[2] D. Xu, Y. Hong, J. Li, and K. Chan, "Effect of tai chi exercise on proprioception of ankle and knee joints in old people," *British journal of sports medicine*, vol. 38, no. 1, pp. 50–54, 2004.

[3] S. Zheng, C. Kim, S. Lal, P. Meier, D. Sibbritt, and C. Zaslawski, "The effects of twelve weeks of tai chi practice on anxiety in stressed but healthy people compared to exercise and wait-list groups–a randomized controlled trial," *Journal of clinical psychology*, vol. 74, no. 1, pp. 83–92, 2018.

[4] S. S.-C. Hui, Y. J. Xie, J. Woo, and T. C.-Y. Kwok, "Effects of tai chi and walking exercises on weight loss, metabolic syndrome parameters, and bone mineral density: a cluster randomized controlled trial," *Evidence-Based Complementary and Alternative Medicine*, vol. 2015, 2015.

[5] K. L. Caldwell, S. M. Bergman, S. R. Collier, N. T. Triplett, R. Quin, J. Bergquist, and C. F. Pieper, "Effects of tai chi chuan on anxiety and sleep quality in young adults: lessons from a randomized controlled feasibility study," *Nature and science of sleep*, vol. 8, p. 305, 2016.

[6] E. Velloso, A. Bulling, and H. Gellersen, "Motionma: motion modelling and analysis by demonstration," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 1309–1318.

[7] F. Anderson, T. Grossman, J. Matejka, and G. Fitzmaurice, "Youmove: enhancing movement training with an augmented reality mirror," in *Proceedings of the 26th annual ACM symposium on User interface software and technology*, 2013, pp. 311–320.

[8] R. Tang, X.-D. Yang, S. Bateman, J. Jorge, and A. Tang, "Physio@ home: Exploring visual guidance and feedback techniques for physiotherapy exercises," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 4123–4132.

[9] M. Sousa, J. Vieira, D. Medeiros, A. Arsenio, and J. Jorge, "Sleevear: Augmented reality for rehabilitation using realtime feedback," in *Proceedings of the 21st international conference on intelligent user interfaces*, 2016, pp. 175–185.

[10] N. Katzakis, J. Tong, O. Ariza, L. Chen, G. Klinker, B. Röder, and F. Steinicke, "Stylo and handifact: Modulating haptic perception through visualizations for posture training in augmented reality," in *Proceedings of the 5th Symposium on Spatial User Interaction*, 2017, pp. 58–67.

[11] A. Ikeda, D.-H. Hwang, H. Koike, G. Bruder, S. Yoshimoto, and S. Cobb, "Ar based self-sports learning system using decayed dynamic timewarping algorithm." in *ICAT-EGVE*, 2018, pp. 171–174.

[12] H.-S. Yeo, H. Koike, and A. Quigley, "Augmented learning for sports using wearable head-worn and wrist-worn devices," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2019, pp. 1578–1580.

[13] Y.-C. Chen, "Pressure sensing insoles for learning tai-chi chuan," in *the Fourth International Conference on Information Science and Cloud Computing*, vol. 264. SISSA Medialab, 2016, p. 048.

[14] J. La Delfa, R. Jarvis, R. A. Khot, and F. Mueller, "Tai chi in the clouds: Using micro uav's to support tai chi practice," in *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, 2018, pp. 513–519.

[15] P. T. Chua, R. Crivella, B. Daly, N. Hu, R. Schaaf, D. Ventura, T. Camill, J. Hodgins, and R. Pausch, "Training for physical tasks in virtual environments: Tai chi," in *IEEE Virtual Reality, 2003. Proceedings.* IEEE, 2003, pp. 87–94.

[16] T. Iwaanaguchi, M. Shinya, S. Nakajima, and M. Shiraishi, "Cyber tai chi-cg-based video materials for tai chi chuan self-study," in *2015 International Conference on Cyberworlds (CW)*. IEEE, 2015, pp. 365–368.

[17] P.-H. Han, Y. Zhong, H.-L. Wang, and Y.-P. Hung, "Augmented visualization for practicing tai-chi chuan with drone-enhanced approach," in *Oz CHI*, 2016.

[18] P.-H. Han, Y.-S. Chen, Y. Zhong, H.-L. Wang, and Y.-P. Hung, "My tai-chi coaches: an augmented-learning tool for practicing tai-chi chuan," in *Proceedings of the 8th Augmented Human International Conference*, 2017, pp. 1–4.

[19] Y.-P. Hung, P.-Y. Kao, Y.-F. Jan, C.-H. Li, C.-H. Chang, and P.-H. Han, "An adaptive tai-chi-chuan ar guiding system based on speed estimation of movement," in *International Computer Symposium*. Springer, 2018, pp. 115–130.

[20] P.-Y. Kao, P.-H. Han, Y.-F. Jan, Z.-W. Yang, C.-H. Li, and Y.-P. Hung, "On learning weight distribution of tai chi chuan using pressure sensing insoles and mr-hmd," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE Computer Society, 2019, pp. 1457–1464.

[21] H. H. Lin, P. H. Han, K. Y. Lu, C. H. Sun, P. Y. Lee, Y. F. Jan, A. M. S. Lee, W. Z. Sun, and Y. P. Hung, "Stillness moves: Exploring body weight-transfer learning in physical training for tai-chi exercise," in *Proceedings of the 1st International Workshop on Multimedia Content Analysis in Sports*, ser. MMSports'18. New York, NY, USA: Association for Computing Machinery, 2018, p. 21–29. [Online]. Available: https://doi.org/10.1145/3265845.3265856

[22] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "Openpose: realtime multi-person 2d pose estimation using part affinity fields," *arXiv preprint arXiv:1812.08008*, 2018.

[23] D. Tome, C. Russell, and L. Agapito, "Lifting from the deep: Convolutional 3d pose estimation from a single image," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2500–2509.

[24] D. S. Alexiadis, P. Kelly, P. Daras, N. E. O'Connor, T. Boubekeur, and M. B. Moussa, "Evaluating a dancer's performance using kinect-based skeleton tracking," in *Proceedings of the 19th ACM international conference on Multimedia*, 2011, pp. 659–662.

[25] Y. Jin, X. Hu, and G. Wu, "A tai chi training system based on fast skeleton matching algorithm," in *European Conference on Computer Vision*. Springer, 2012, pp. 667–670.