A Framework for Correcting Human Motion Alignment for Traditional Dance Training Using Augmented Reality

Javid Iqbal, and Manjit Singh Sidhu
Universiti Tenaga Nasional, Malaysia, manjavid@gmail.com

ABSTRACT
This paper presents a framework for motion capture analysis for dance learning technology using Microsoft Kinect V2. The proposed technology utilizes motion detection, emotion analysis, coordination analysis and interactive feedback techniques for a particular dance style selected by the trainee. This motion capture system solves the heterogeneity of the existing dance learning system and hence provides robustness. The analysis of the proposed work is carried out using query techniques and heuristic evaluation. The Microsoft Kinect V2 embedded with Augmented Reality (AR) technology is explored to demonstrate the recognition accuracy of the proposed framework.

Keywords: Microsoft Kinect V2, Augmented Reality, Dance, learning technology.

I INTRODUCTION

Dance is an ethereal art since it depends on the motion of performers gestures. Dance can communicate various messages according to the context, and focus on aesthetical aspects (classical, contemporary, western dance) traditional aspects (folk, cultural dances) and spiritual aspects and so on. The cultural dances are strongly connected to the traditional heritage that shapes the identity of a particular place. These dance styles are explored and learnt through small group of people who get together to practice and gain expertise on these traditional dances. Hence, know- how aspects of these dance forms face a major risk factor where few elements of this intangible cultural heritage could die out if not preserved and safeguarded for the future generations.

Information technology plays a vital role for development of platforms that are employed for analysis, capturing and modeling of such extensive dance interactions and can eventually contribute towards significant transformation of artistic knowledge for next generation. The main drawback lies in the precise recognition of human body movements. Nowadays, the evolution of motion capture technology that has led to the innovation of Microsoft Kinect sensors have advantages over earlier systems to make more precise measurements with wide array of sensing capabilities, availability of processing power to achieve complex data interpretations and enhance the flexibility of dance learning technology.

A. Dance Technology

The advancement in technology has paved way for easy dance education. Computer Vision (CV) and Augmented Reality (AR) are the resent trends in dance education. There are more than 100 different dance styles in world. Is it possible for an individual to learn different dance styles? Yes, it is possible but the learner must find different dance professionals or choreographers to learn different dance styles. The dance learner must make sure he goes to different dance schools to learn different dance styles, it is also mandatory that the learner can learn the specific dance style from the professional only during scheduled time.

AR technology steps in to overcome all the limitations in traditional dance learning. Technology has made available numerous ways to search for information or video to make self learning easier, but the major drawback is that the online websites do not provide interactive feedback and inspiration to the learners by which the learning process becomes unprofessional.

To learn a subject we need books to learn a specific dance style we need a professional dancer or choreographers. We propose a system using Kinect V2 which bring in the choreographer for in house and class room training with interactive feedback and comments.

II RELATED WORK

Computer vision and motion sensing technology have enabled the users to actively, physically and mechanically interact with the digital environment in various ways. The hybrid combinations of traditional art forms and advanced CV technology have made the authors in the last decade to drive out the AR based dance learning system. Researches provide an updated interactive performance system for floor and aerial dance that controls visual and sonic aspects of the presentation through Microsoft Kinect camera. Improvised gesture recognition and tracking system called Action Graph (AG) is described in this article, which has the capability of capturing incoming gestures in an unsupervised way & enables mapping between input gestures to desired rendering functionalities (Wang, 2015). In this paper, authors have presented five interactions used in augmenting an improvised dance show. The system that allows those interactions is composed of several interconnected modules (Clay et al., 2012). The researchers have described Cha learn gesture data set that is user dependent, small vocabulary and one-shot learning using Kinect camera (Guyon et al., 2014). Authors present a simple and computationally efficient framework for 3D dance basic motion recognition based on syntactic pattern recognition (Heryadi et al., 2013). In this paper, authors measure the desirability of the tool for describing and reproducing Thai dance. The survey experiment was conducted on over 200 students and teachers in four well-known schools in Thai dances. The results show that about 70% of subjects think that the software has usability, desirability, creativity and fun, (Choensawat et al., 2013). The following Table 1 illustrates the research findings and drawbacks of the existing systems in AR based dance learning technology.

Table 1. Summary of drawbacks and research findings for AR based dance learning technology

Year	Author	Techniques / methods used	Drawbacks	Research findings
2015	Dubnov and Wang	MS Kinect, Action Graph (AG), Multiple IR Markers.	The authors have not focused on errors and feedback.	There is no consideration of interactive feedback and more number of trainees.
2014	Guyon et al.	Small vocabulary, fixed camera, one-shot learning, and data include 54,000 hand and arm gestures recorded with an RGB-D Kinect camera	Different gestures are learnt no dance is considered and no feedback is taken into consideration.	Different dance styles are necessary to master dance.
2013	Heryadi, Fanany, and Arymurthy	Syntactic pattern recognition, Stochastic Regular Grammar (SRG)	Dance basic motion analysis is done further work must be done for dance motion analysis; No feedback is given	Only body motions are considered. Dance motion analysis is to be examined.
2013	Choensawat, Sookhanaphibarn, and Kijkhun	Labanotation, LabanEditor, Thai Dance Notation System.	Authors focus only on Thai dance style. No feedback or comments given from the system	Only one dance style is focused.
2013	Anderson et al. 2013	Half-silvered mirrors (HSMs), Microsoft Kinect	The Kinect has difficulty tracking movements that cause large amounts of occlusions; Improvements in sensing technology will open new domains of training.	Recorded data is used for training the users and problems faced by large amount of occlusion.
2013	Kuramoto et al.	Augmented Practice Mirror (APM), Radish the motion tracking system	Authors focus on velocity and acceleration visualization method for APM, for learner's further understanding of teacher's motion.	Other parameters such as error and feedback are to be considered.

The authors have discussed a novel system YouMove that allows users to record and learn physical movement sequences. The Kinect-based recording system is designed to be simple, allowing anyone to create and share training content, some of the screen shots are shown in Figure 1. The corresponding training system uses recorded data to train the user using a large-scale AR mirror. The system trains the user through a series of stages that gradually reduce the user's reliance on guidance and feedback. This also discusses the design and implementation of YouMove and its interactive mirror. The authors have presented a user study in which YouMove was shown to improve learning and short-term retention by a factor of 2 compared to a traditional video demonstration. While the presented implementation uses a half-silvered mirror as a display, the software could also run as a traditional video-based AR system (Anderson et al., 2013). The Kinect has difficulty tracking movements that cause large amounts of occlusions. This would be more accessible to users, but does not provide the real-time feedback that the mirror does. It would be interesting to better understand any learning difference between a mirror and video based system on various devices (large screen, small screen, etc.). The addition of social features and richer inclusions of gaming technologies could also greatly help YouMove. The paper presents an algorithm for real-time body motion analysis for dance pattern recognition by use of a dynamic stereo vision sensor and Hidden Marcovian Method (HMM) (Kohn et al., 2012).

Authors had illustrated a framework to develop a digital bharathanatiyam interaction (Majumdar, 2012). A novel approach is proposed for generating dance performance based on music similarity (Lee et al., 2012). The research had presented an automatic dance lesson generation system which is suitable in a learning-by-mimicking scenario where the learning objects can be represented as multi-attribute time series data (Yang et al., 2012). Authors have described a novel framework for music-driven dance choreography synthesis and animation. For this purpose, authors construct a many-to-many statistical mapping from musical measures to dance figures based on the correlations between dance figures and musical measures as well as the correlations between successive dance figures in terms of figure-to-figure transition probabilities. Then use this mapping to synthesize a music-driven sequence of dance figure labels via a constraint based dynamic programming procedure. With the help of exchangeable figures notion, the proposed framework is able to yield a variety of different dance figure sequences (Ofli et al., 2012).

This paper has provided an overview of the contextual issues that surrounds the design of two interactive applications for building awareness of ICH in museums. The agenda is to design a serious gaming environment for visitors in which to learn indigenous dance inside the museum (Khan and de Byl, 2012). This study makes a body model for education of dance how lively a dance trainer moves his/her body so that the dance student can learn dance from the dance trainer who is in the distant place through communication line in virtual dance studio (Takai, 2012). The authors have proposed a new dance training system based on motion capture and virtual reality technologies (Chan et al., 2011). Authors proposed a visualization method of the velocity and acceleration of the teacher's motion for the learner to understand it more easily and Clearly (Kuramoto et al., 2013). The research analysis done for dance learning with or without interactive feedback shows that there is no consideration for more number of users in literature till date as depicted in Table 2.

Table 2. Summarization of dance learning technology

Year	Author	Augmented (Yes/ No)	Type of art	Interactive feedback	No of trainees
2015	Dubnov and Wang	Yes	Floor and Aerial Dance	No	1
2014	Gnyon et al.	No	Gesture	No	1
2013	Heryadi, Fanany, and Arymurthy.	No	Basic Dance	No	1
2013	Choensawat, Sookhanaphibarn, and Kijklum.	No	Thai Dance	No	1
2013	Anderson et al.	Yes	Dance	Yes	1
2013	Kuramoto et al.	Yes	Physical motion	No	
2012	Kohn, Nowakowska, and Belbachir.	No	Dance pattern	No	1
2012	Majumdar.	No	Bharatanatayan dance	No	1
2012	Lee, Lee, and Park.	No	Dance	No	1
2012	Yang et al.	No	Dance	No	1
2012	Khan and de Byl.	Yes	Indigenous dance	No	1
2012	Takai.	No	dance	No	1
2012	Clay et al.	Yes	Ballet Dance	Yes	1
2012	Offi et al.	No	Dance	No	1
2011	Chan et al.	No	Dance	Yes	1

Figure 1. AR based dance learning using Kinect

III SYSTEM DESIGN

The agenda of the research is to use full body interaction for two to six numbers of trainees with interactive feedback. The hardware and the software of the system are comprises of Microsoft Kinect V2 which generates the skeletal image of the trainee who stands in front of Kinect V2. Kinect V2 has the maximum capabilities of tracking six numbers of trainees with 26 body joints per trainee. Kinect V2 supports USB 3.0 since the number of process done by the kinect must be transferred to a system for processing with optimum transfer speed which is obtained only with USB 3.0 and not with USB 2.0. The architecture of the system design/ Kinect SDK is shown in Figure 2.

Kinect for Windows can work with windows 8.1 and windows 10. Microsoft visual studio is used for developing web application and web services using coding languages like C# and XAML codes for Kinect. To use the audio control that is to get the interactive feedback Kinect audio control SDK must be installed.

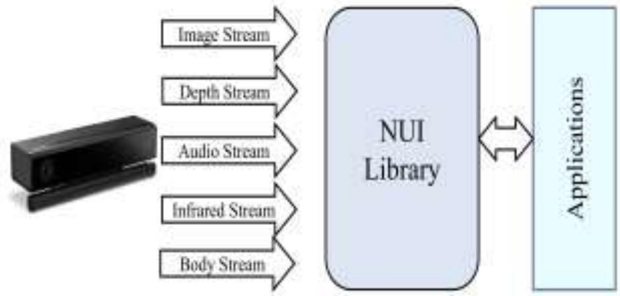

Figure 2. System design

A. Interface/ Model

The interface of the system is designed in such a way that even a beginner can interact with the system. The hand gesture reorganization in Kinect v2 is used as interaction platform for the trainee and the trainers. With the use of hand gesture reorganization the trainee need not move from their location to the system to interact. The trainee must raise their hands to start the interaction with the system, once the hand is recognized by the system the hand icon will appear on the screen after which the trainee can scroll, click and select the option available on the screen with his hands. A screen shot of the hand gesture reorganization is given in Figure 3.

B. Interaction

There are two interactions involved in the system they are Trainee interacting with the system and the system giving feedback to the trainee.

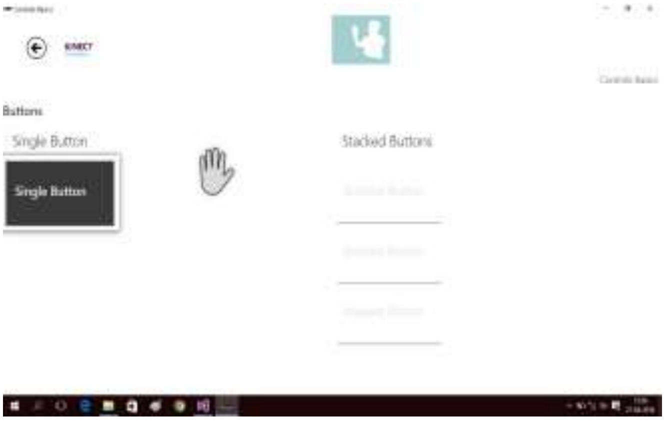

Figure 3. Hand gesture recognition.

Trainee interacting with the system: The trainee interacting with the system is uses friendly that even a beginner can interact with the menus and options available. Some of the menus available are different types of dance styles available, our research mainly focuses on different dance styles available in India, some of the famous dance styles are Bharathanatiyam, kathakali, Oddisi, Bhangra and many more.

The trainee can choose any dance forms available in the main screen by clicking on a specific dance style using hand gesture reorganization the menu jumps into the selected submenu where the trainee will have a brief description about the selected dance and what level of dance he/ she wants to learn with the system.

Different levels of dance are Beginner, Intermediate and Advanced. There are three different levels of dances available for the learner to choose from according to the selected dance level the training starts.

System giving feedback to the trainee: After selecting the dance style the trainee can choose different steps available in the database. Once the level and the steps are selected, the skeletal image of the selected dance will be projected on the screen the trainee who wants to learn dance must match the skeleton movement on the screen with the skeletal of their own. If there is 80 % match in the movement then the next step is continued. If there is a mismatch in the steps performed then the system will indicate the error with a balloon on the top of the trainee skeleton stating there is a miss match in the steps kindly move your leg to the skeleton on the screen and also a voice response telling the same to the trainee. Once the step is attained 80% accuracy there will be a motivating feedback from the system "Excellent we shall move on to the next step". Thus the proposed system helps the trainee with interactive feedback and motivating comments.

IV PROPOSED FRAMEWORK

The architecture of the proposed system consists of three major components training using Kinect, Motion capture and performance evaluation. Figure 4 shows the relation between each of the components. The students dance movements are captured by Kinect V2 and are detected for subsequent correctness of the motion using the motion detector. The motion detector uses motion analysis method and motion capture technology. The emotion expressed by the dancers are analyzed and matched with the data in the data base through the emotion analyzer using emotion analysis technology. The interactive feedback component provides feedback and comments to the student. The comments and the feedback given to the trainers are displayed using a balloon text on the screen and a voice reading the same. The performance evaluation compares the movements performed by the student based on the repetition and provides individual scores for every dance step performed. The computing devices are used to visualize the overall movements done by the student by a physical teacher either at the same place or from a remote location.

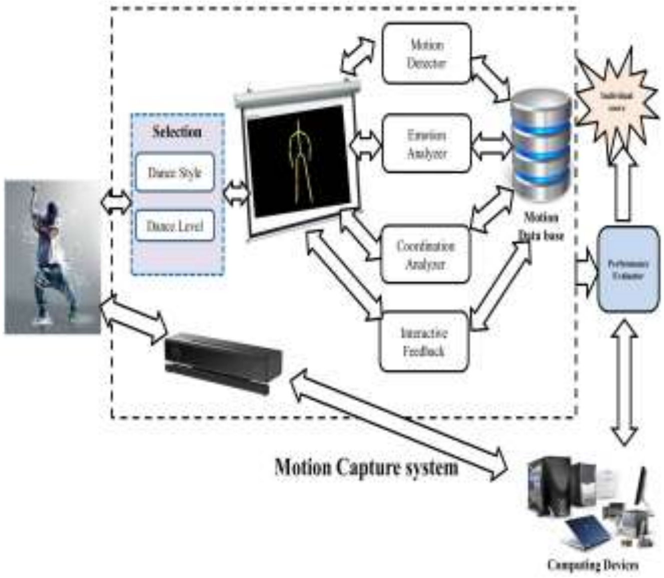

Figure 4. Proposed Framework

V ANALYSIS METHODOLOGY

The analysis of this research deals with two major aspects firstly the design of the system is evaluated followed by evaluation through user participation and query techniques. The trainees of the proposed dance learning system are divided into two main groups namely control group and study group. Control group is defined as a group of trainees who are expertise in dance skills with the help of trainers, whereas study group consists of trainees who are trained with kinect V2. Each of these groups consists of 5 batches where each batch has 6 numbers of students. The different criteria upon which design analysis and usability test are done includes are group, dance style/type, place/ country users opinion about learning dance / physical activity and users activity of interest.

The evaluation of system design is done using cognitive walkthrough method, heuristic evaluation and review based evaluation. The study in control group and study group are trained intensively for one month and then the mode of training is interchanged between the groups to evaluate the impact of training by dance experts.

In this way the performance, repeatability precision and efficiency of the dance is evaluated. Evaluation through user's participation is done based upon queries techniques such as interview and survey. The users those who have expertise their dance skills based upon the proposed dance learning system are interviewed and a detailed survey is conducted as how well they have been trained. The performance analysis of control group and study group is done using correlation analysis to gain reliability of the proposed system and the results will be published in the next subsequent paper.

VI **CONCLUSION**

The significant contribution of AR towards learning technology in computer vision has led many researchers to adopt AR as one of the most promising and distinctive direction for computer vision education. The users can experience the effect of individual in- house training as well as group learning for personal motivation and enhancement of artistic skills. This paper describes a framework for dance training with interactive feedback. The advantage of AR in educational sector paves way to visualize digital media, in- depth observation of subjects involved and possibilities to assess the virtual information according to the users need. The evolution of Microsoft Kinect makes AR based dance learning more fun- filled, user-friendly, self-motivating and responsive experience. It can further be concluded that traditional and cultural dance forms can be preserved and safeguarded to transform it to the future generations with the help of CV technology.

REFERENCES

Anderson, Fraser, Tovi Grossman, Justin Matejka, and George Fitzmaurice. (2013). "YouMove: Enhancing Movement Training with an Augmented Reality Mirror." *Proceedings of the 26th annual ACM symposium on User interface software and technology - UIST '13*: 311–20. http://dl.acm.org/citation.cfm?id=2501988.2502045.

Chan, J C P, H Leung, J K T Tang, and T Komura. (2011). "A Virtual Reality Dance Training System Using Motion Capture Technology." *IEEE Transactions on Learning Technologies* 4(2): 187–95.

Choensawat, Worawat, Kingkarn Sookhanaphibarn, and Chommanad Kijkhun. (2013). "Desirability of a Teaching and Learning Tool for Thai Dance Body Motion." : 171–79.

Clay, Alexis et al. (2012). "Interactions and Systems for Augmenting a Live Dance Performance." *11th IEEE International Symposium on Mixed and Augmented Reality 2012 - Arts, Media, and Humanities Papers, ISMAR-AMH 2012*: 29–38.

Fan, Rukun, Songhua Xu, and Weidong Geng. (2012). "Example-Based Automatic Music-Driven Conventional Dance Motion Synthesis." *IEEE transactions on visualization and computer graphics* 18(3): 501–15. http://www.ncbi.nlm.nih.gov/pubmed/21519104.

Guyon, Isabelle, Vassilis Athitsos, Pat Jangyodsuk, and Hugo Jair Escalante. (2014). "The ChaLearn Gesture Dataset (CGD 2011)." *Machine Vision and Applications* (Cgd): 1–23.

Heryadi, Y, MI Fanany, and AM Arymurthy. (2013). "Stochastic Regular Grammar-Based Learning for Basic Dance Motion Recognition." *Cloud2.Snappages.Com*: 978–79. http://cloud2.snappages.com/b887f6216920e33921eb3600dca196648e63c23d/767-SRG for Dance Basic Dance Motion Recognition.pdf.

Khan, Muqeem, and Penny de Byl. (2012). "Creating Tangible Cultural Learning Opportunities for Indigenous Dance with Motion Detecting Technologies." *2012 IEEE International Games Innovation Conference*: 1–3. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6329834.

Kohn, Bernhard, Aneta Nowakowska, and Ahmed Nabil Belbachir. (2012). "Real-Time Body Motion Analysis For Dance Pattern Recognition." *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*: 48–53. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6238894.

Kuramoto, Itaru et al. (2013). "Visualizing Velocity and Acceleration on Augmented Practice Mirror Self-Learning Support System of Physical Motion." *2013 Second IIAI International Conference on Advanced Applied Informatics*: 365–68. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6630375.

Lee, Minho, Kyogu Lee, and Jaeheung Park. (2012). "Music Similarity-Based Approach to Generating Dance Motion Sequence." *Multimedia Tools and Applications* 62(3): 895–912. http://link.springer.com/10.1007/s11042-012-1288-5.

Majumdar, Rwitajit. (2012). "Framework for Teaching Bharatanatyam through Digital Medium." : 241–42.

Ofli, Ferda, Engin Erzin, Yücel Yemez, and a. Murat Tekalp. (2012). "Learn2Dance: Learning Statistical Music-to-Dance Mappings for Choreography Synthesis." *IEEE Transactions on Multimedia* 14(3 PART 2): 747–59.

Takai, Miwa. (2012). "Production of Body Model for Education of Dance by Measurement Active Quantity." *The 1st IEEE Global Conference on Consumer Electronics 2012*: 212–16. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6379584.

Wang, Cheng-i. (2015). "Free-Body Gesture Tracking and Augmented Reality Improvisation for Floor and Aerial Dance Free-Body Gesture Tracking and Augmented Reality." (October).

Yang, Yang, Howard Leung, Lihua Yue, and Liqun Deng. (2012). "Automatic Dance Lesson Generation." *IEEE Transactions on Learning Technologies* 5(3): 191–98.