



ISSN:1991-8178

## Australian Journal of Basic and Applied Sciences

Journal home page: www.ajbasweb.com



### A System Framework for Computer Vision and AR based Dance Learning Technology

<sup>1</sup>Javid Iqbal and <sup>2</sup>Manjit Singh Sidhu

<sup>1,2</sup>Universiti Tenaga Nasional, Department of Graphics and Multimedia, College of IT, 43000 Kajang, Malaysia.

#### ARTICLE INFO

##### Article history:

Received 12 March 2015

Accepted 28 April 2015

Available online 2 May 2015

##### Keywords:

Augmented Reality, Computer vision, Multimedia technology, Dance Learning technology, Kinect V2

#### ABSTRACT

**Background:** Computer vision (CV) has offered various benefits for Augmented Reality (AR) technologies in the field of learning, construction, manufacturing, navigation, information modelling, visualization and communication. In the past few years many choreographers have focused upon implementation of computer technology to enhance their artistic skills. CV technology presents new methods for learning, instructing, developing and assessing physical movements as well as provides scope to expand dance resources and rediscover the learning process. This paper presents a system framework for interactive dance training using kinect V2. The authors also aim to provide an overview for learning standards based on AR Kinect sensors. In addition the future work is towards exploring the latest version of Kinect V2 for dance training that could become the next futuristic virtual educator.

© 2015 AENSI Publisher All rights reserved.

**To Cite This Article:** Javid Iqbal and Manjit Singh Sidhu., A System Framework for Computer Vision and AR based Dance Learning Technology. *Aust. J. Basic & Appl. Sci.*, 9(19): 27-34, 2015

#### INTRODUCTION

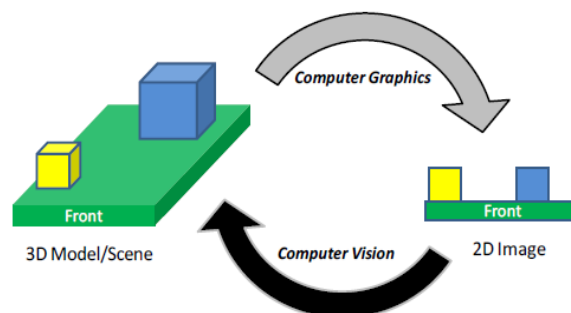
##### Overview of Ar and Computer Vision:

In general, the increase in computing performance has enabled many fields, such as AR and computer vision, to advance over the years. AR is defined as a technology where real world objects can be seen superimposed by virtual machine generated entities in order to provide interaction between real world and virtual world objects. AR has been employed in numerous areas including construction, automobile, aerospace, military, medicine and education to bridge the gap between visualization and information retrieval. In recent years, development of human-computer interaction has improved significantly. AR acts a connecting bridge by eventually combining the real world and virtually generated objects together and allows the user to view it seamlessly in the same environment. This enhances the usability and information retrieval which human beings cannot perceive with their regular senses.

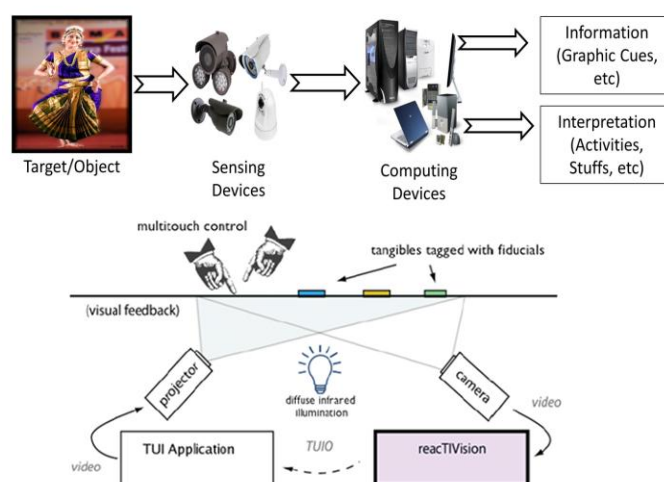
Many computer vision applications have a graphics component, such as rendering the results of 3D reconstruction. This has led to CV being typically linked with computer graphics and the two together are referred to as visual computing (Randy, 2001) However, a better summation of the relationship between them is computer vision as the inverse of

computer graphics. In computer graphics, an application has a 3D model of the world and displays this information by projecting onto the 2D plane of the screen. In computer vision the system is given information about the projection in the form of image data and attempts to extract information about the 3D world that created the projections. (Fig. 1) graphically depicts the inverse relationship these two field share.

The typical CV system involves three components, as shown in (Fig. 2) (left image). The first is the scene under study. The second is a sensing device that can be used to analyze the scene. The third is a computational device that can perform the analysis of the scene based on the data from the sensor. The computation device generates two possible forms of data, information such as visual cues, and interpretations of information such as actions being performed or the presence of objects. The two forms of data can each be used to refine the other until the output of the vision system is computed with a predefined amount of certainty. The result can be the 3D location for every pixel within and image or the certainty that a person is performing jumping jacks. With proper selection of the information and the interpretation algorithm, CV systems can be applied in a large number of applications. The architecture of CV is depicted in (Fig. 2) (right image) (Randy, 2001).



**Fig. 1:** Inverse Relationship of Computer Vision and Computer Graphics. Computer vision maintains a relationship to the field of computer graphics. In computer graphics the system has a 3D model as seen on the left and computes how to project that 3D scene onto the 2D plane of a display as seen on the right. In computer vision the problem begins with the 2D information on the right and attempts to reconstruct the 3D model and/or context on the left.



**Fig. 2:** Typical Computer Vision System (left image) and architecture (right image). The figure shows the basic components of a computer vision system. The left most component is the scene or object of study. The next required component shown in the figure is the sensing device used to collect data from the scene. The third component is the computation device. The device computes information such as visual cues and reasons on this information to generate interpretations of the scene such as objects present or actions being performed.

### **Dance Learning Technology:**

Learning dance and mastering in a particular dance style can turn out to be the most challenging task for new learners. Although this will need a choreographer or professional training sessions, technology has replaced such traditional way of learning. People who consider dance as their hobby or who hardly find time to look for a trained teacher or attend any dance classes; will mostly go for self-learning or in-house practice. 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 any interactive feedback and inspiration to the learner by which the learning process becomes less interesting and ineffective.

### **Background Study:**

#### **Ar In Education And Training:**

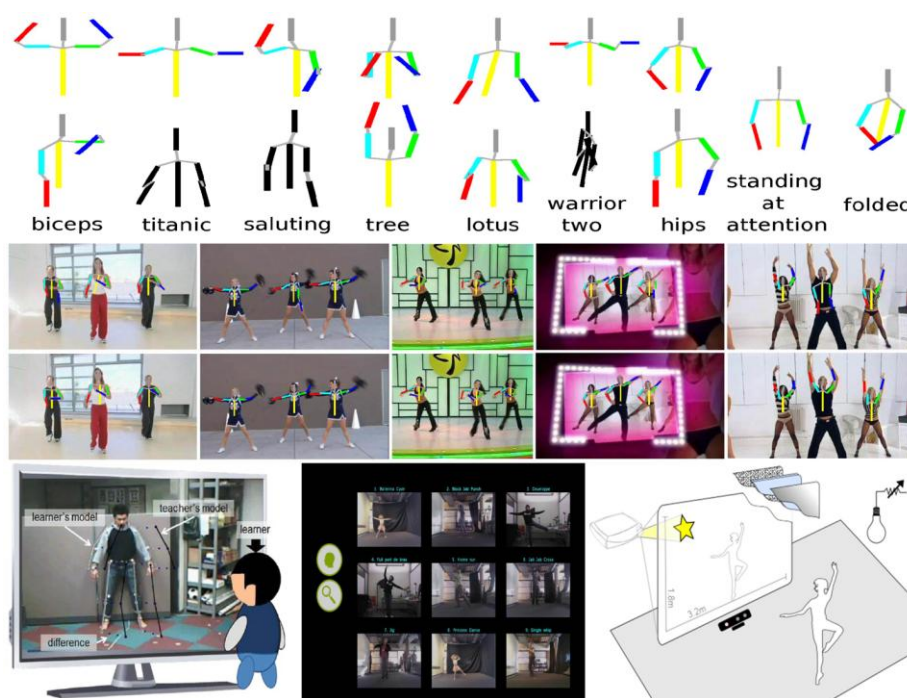
Augmented reality as defined earlier is a technology that makes use of CV techniques to collaborate computer generated virtual objects with real time environment in order to increase or to enhance what can be visualized by the human user (Hall, 2001). In educational sector AR can be used in dance education, Chemistry, Biology, Astronomy, Mathematics, Computer Graphics etc. AR in chemistry education is used for the exploration of physical models of amino acids (Fjeld, 2002). In biology AR based learning system is used to get insights into the interior of human organs on a detailed basis with easier understanding (Gillet, 2004). In astronomy AR technology is applied for better understanding of seasonal changes in light and temperature, rotation/revolution of Sun and Earth (Brett, 2002). AR is also used as a visualization tool in computer graphics laboratories and in computer aided design lectures (Kaufmann, 2008 and Chen, 2006). In mathematics AR is used for teaching

calculus and algorithms Kaufmann, (2002). In the field of dance education Kinect sensors are employed to train the students in physical movements as well as master them in their skills.

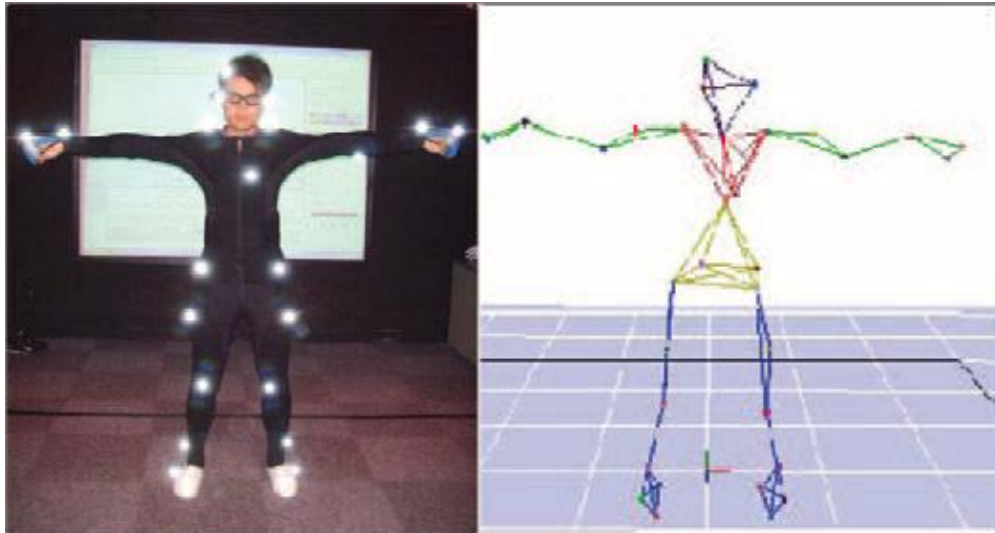
### **Fun Filled Ar Based Dance Learning Using Kinect:**

Computer vision and motion sensing technologies have enabled the users to actively, physically and mechanically interact with the digital environment in varied ways. The hybrid combinations of traditional art forms and advanced CV techniques have made the authors in the last decade to drive out to AR based dance leaning systems. Guyon, (2013) describes Cha learn gesture data set that is user dependent, small vocabulary and one-shot learning using Kinect camera. Silva, (2013) developed a prototype to monitor fall risk while playing a game using smart phone accelerometer. Fan, (2012) introduced a novel method for synthesizing dance motion that follows the emotions and the contents of a piece of music. Yang, (2012) 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. Chan, (2011) proposed a new dance training system based on motion capture and virtual reality technologies. Golshani, (2004) presented multimedia information repository for cross cultural dance studies such as East Indian dance with the use of two 3D Vicon motion capture systems. Eichner, (2012) proposed techniques for novel Human Pose Co-estimation for joint pose estimation over multiple persons in a common, but unknown, pose. Kuramoto, (2013) a

visualization method of velocity and acceleration of teacher's motion for the learner to understand more clearly and easily has been proposed. Anderson, (2013) 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 (Fig. 3). 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. 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. One can imagine online yoga, dance or martial arts classes, with competition from online peer groups, but more work is needed to achieve this.



**Fig. 3:** Fun filled AR based learning using Kinect



**Fig. 4:** The choreographer with optical marker suite a (left image) and the mesh model (right image)

The motion data as explained by Yang, 2012 captured by a marker-based optical 3D motion capture system with several cameras. The individual shown in the (Fig. 4) (right image) wears a sensor suit attached with 3D optical markers on different body parts. The motion capture methodology is

capable of tracking and recording human motion. The 3D motion of the actor is captured as a time series data where each frames contains the coordinates of the 3D marker. A mesh model is shown in the right image of the Fig.4 can be fit on the skeleton to generate a 3D virtual model.



**Fig. 5:** Layout of 3D Viewer (left image) and the real movement by a dancer (right image)

Chan, 2011 had presented in his work a virtual teacher demonstrating a dance motion as depicted in (Fig.5) left image and the right image shows that actual movement performed by a real dancer. This solution makes use of motion capture technology and motion analysis method. The virtual teacher also appears when the student is practicing the moves, thus enabling the student to imitate the moves of the

teacher. Eventually the students may also be able to notice the error of timing and performance of the movement at a glance. In order to facilitate the observation the authors also apply a mirroring effect to the virtual agents in the screen such that the student can learn under the same setup as in a dance class room.

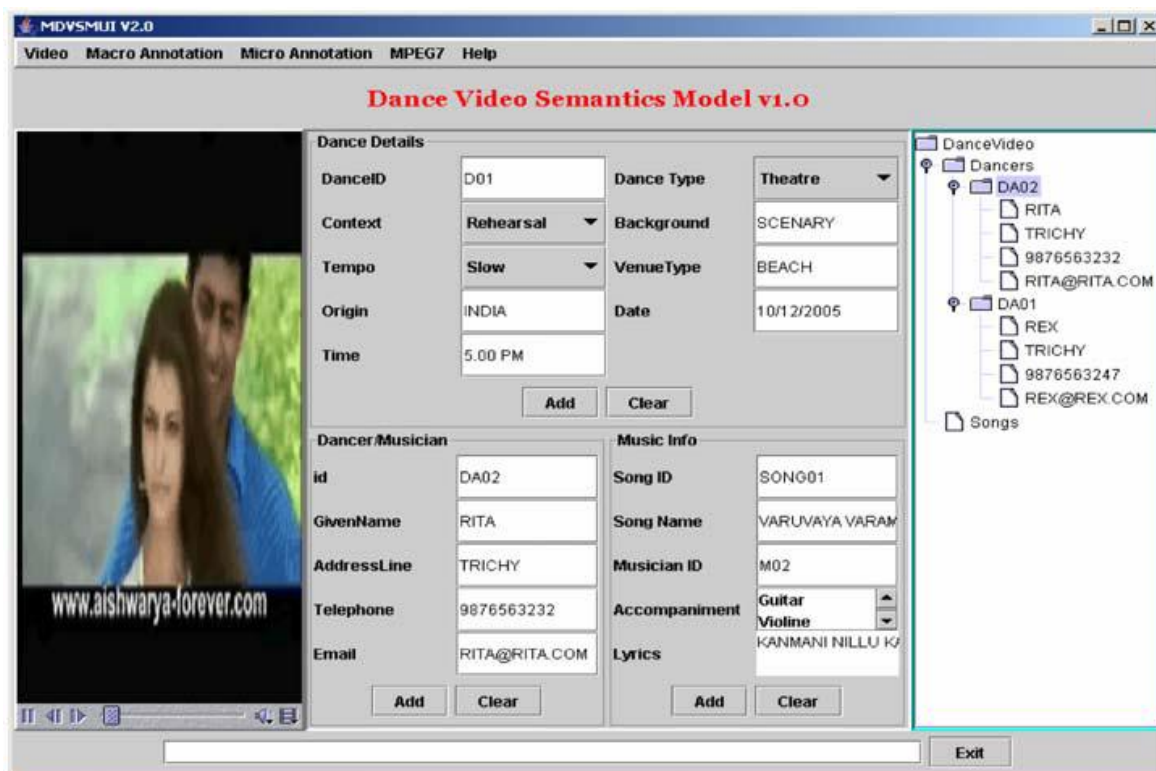


Fig. 6: Screenshot of macro annotation

Kannan, (2010) had presented a system named DanVideo had been presented for semi-automatic authoring and access to dance archives. DanVideo provides methods of annotation and authoring and retrieval tools for choreographers, dancers and students. The authors have demonstrated how dance media can be semantically annotated and how this information can be used for the retrieval of dance video semantics. The screen shot in Fig. 6 depicts the rendering of the dance for a song in the Bollywood movie “*Hai Mera Dil*” and the macro annotator. The

macro annotator describes the macro features. The set of semantics that are annotated include the details of the dancers (such as ID, GivenName, PostalAddress, Telephone and Email), details of the musician (similar to those of dancers), music, song, background, tempo of the dance steps (slow, medium, or fast), dance origin, dance type, context (live, rehearsal, professional play, competition, etc.), date and time of recording, and type of performance venue (theatre, beach, etc.).

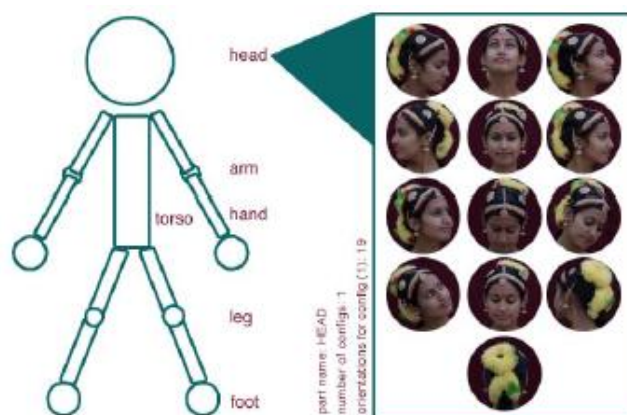


Fig. 7: Body parts classification & library elements for head

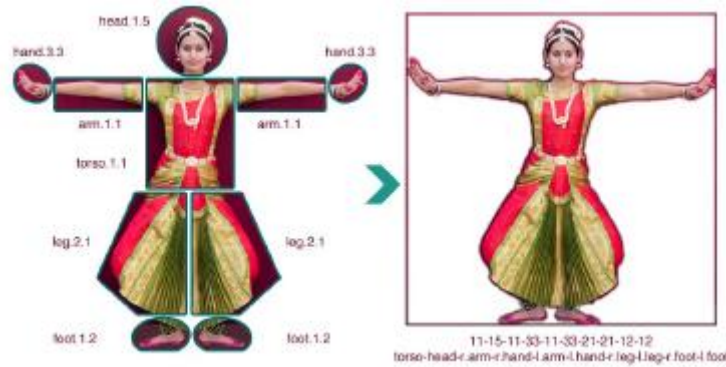


Fig. 8: Body posture generated by assembly of body parts

Majumdar, (2012) had illustrated an effort to develop a digital Bharatanatyam interaction. A hierarchical architecture is extracted from the dance movements by analyzing its grammar which is then used in developing a framework to digitize the dance forms. The Indian dance, like an Indian sculptor, does not place much emphasis on the muscles of the human body but takes the joints and the fundamental bone structure as its basis. Based on this the bottom up approach can be taken to assimilate a dance piece. The body is divided into parts each having possible configurations that appear in dance. Corresponding orientations of each configuration are counted. Indexing of all possibilities is done as Part, Configuration and Orientation. This generates the library of body parts with all the possible

configurations and orientation. (Fig. 7) describes the schematics of division of body and illustrates an example of library elements for the head. Elements are selected from each body part library and are assembled to define a body posture. Assemblies of kinesthetically valid and correct postures are governed by the first set of rules. Each posture is specified as an ordered list, defined by the index number of the individual parts. (Fig.8) illustrates the concept of assembly of individual body parts into a body posture. Thus at this stage a second library is generated that keeps the list of all possible body postures.

**System Overview:**

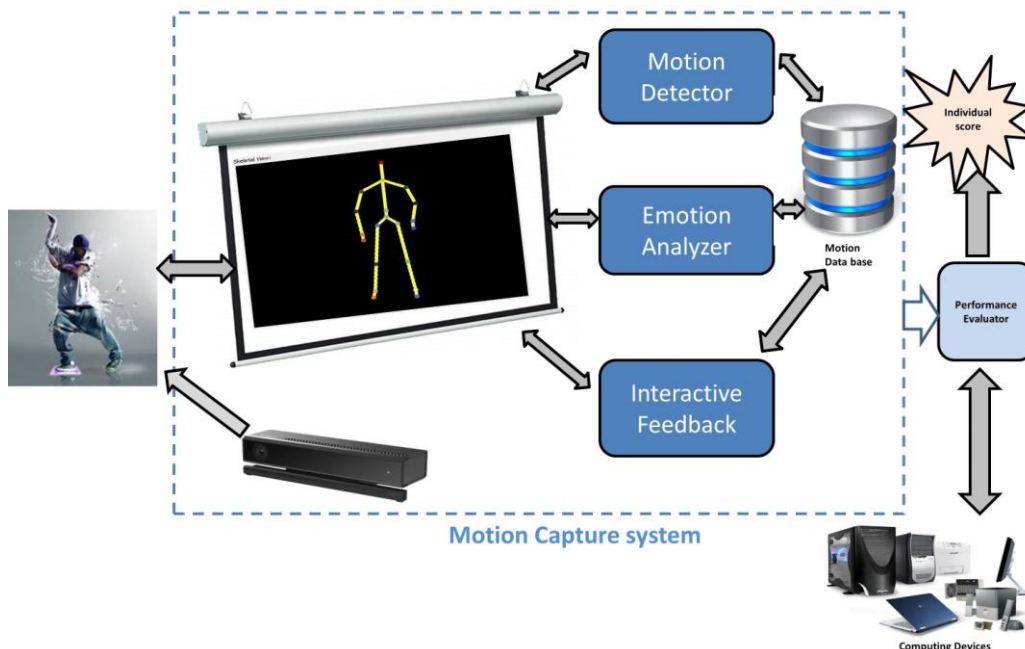


Fig. 9: Proposed Architecture

The architecture of the proposed system consists of three major components training using Kinect, Motion capture and performance evaluation. (Fig.9) 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 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.

#### ***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 users 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.

#### ***Conclusion:***

Augmented reality has made distinctive contributions towards learning experiences. The developments in CV technology have led researchers to enhance and assess AR learning aspects. The evolution of learning paradigm from traditional methods towards CV, relative to the theory of education technology has showed a significant transformation and has eventually led researchers and technologists to adopt AR as one of the promising direction for multimedia learning

technology. The insights for AR technology in navigation and construction industries provide an essential need for comprehensive systems, integration of multiple platforms, user friendly interfaces, defect detection and seamless integration at an affordable cost. The new proposed architecture presented in this paper would be useful in designing a futuristic digital choreographer with feedback processing capability and help many enthusiastic dances to learn how to dance. The AR technology can be used for group as well as individual learning which motivates the learner in every possible way. This paper provides insights into the proposed interactive dance learning methodology to expertise oneself in dance or any other physical activities. The advantage of AR in educational sector includes opportunity to visualize digital information, observe the finer details of subjects and possibilities to examine the virtual information perceptively as many times as needed. There has also been emergence of AR based dance learning technology with tracking capabilities such as Microsoft Kinect. This technology has been employed for fun-filled dance learning with responsive interactions and self-motivating feedback. It can further be concluded that the traditional dances can be explored and the cultural heritage can be preserved with CV technology.

#### **REFERENCES**

- Anderson, F., T. Grossman, J. Matejka and G. Fitzmaurice, 2013. YouMove: Enhancing Movement Training with an Augmented Reality Mirror (pp.311-320), 26th Annual ACM Symposium on User Interface Software and Technology.
- Brett, E.S and R.H. Nicholas, 2002. Using Augmented Reality for Teaching Earth-Sun Relationships to Undergraduate Geography Students, IEEE International Augmented Reality Toolkit Workshop.
- Brett, E.S., 2002. Augmented Reality and Education- Current Projects and the Potential for Classroom Learning, Journal on New Horizons for Learning.
- Bruin, E.D., D. Schoene, G. Pichierri, S.T. Smith, 2010. Use of virtual reality technique for the training of motor control in the elderly, Springer journal of ZeitschriftfurGerontologie und Geriatrie.
- 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.
- Chen, Y.C., 2006. A study of comparing the use of augmented reality and physical models in chemistry education, Proceedings of ACM international conference on Virtual reality continuum and its applications.

Eichner, M., and V. Ferrari, 2012. Human Pose Co-Estimation and Applications, IEEE transactions on Pattern Analysis and Machine Intelligence.

Fan, R., S. Xu and W. Geng, 2012. Example-Based Automatic Music-Driven Conventional Dance Motion Synthesis, IEEE transactions On Visualization and Computer Graphics.

Fjeld, M. and M.V. Benedikt, 2002. Augmented Chemistry: An Interactive Educational Workbench, IEEE International Symposium on Mixed and Augmented Reality.

Geroimenko, V., 2012. Augmented Reality Technology and Art: The Analysis and Visualization of Evolving Conceptual Models, International Conference on Information Visualisation.

Gillet, A., M. Sanner, D. Stoffler, D. Goodsell and A. Olson, 2004. Augmented Reality with Tangible Auto-Fabricated Models for Molecular Biology Applications, ACM International Conference on Visualization.

Golshani, F., P. Vissicaro and Y. Park, 2004. A Multimedia Information Repository for Cross Cultural Dance Studies (pp.89–103), Multimedia Tools and Applications.

Guyon, I., V. Athitsos, P. Jangyodsuk, H.J. Escalante and B. Hamner, 2013. Results and Analysis of the ChaLearn Gesture Challenge 2012 (pp. 186–204), Springer book on Advances in Depth Image Analysis and Applications.

Guyon, I., V. Athitsos, P. Jangyodsuk and H.J. Escalante, 2014. The ChaLearn gesture dataset (CGD 2011), Springer Journal of Machine Vision and Applications.

Hall, T., L. Ciolfi, L. Bannon, M. Fraser, S. Benford, J. Bowers, C. Greenhalgh, S.O. Hellstrom, S. Izadi,

H. Schnadelbach and M. Flintham, 2001. The Visitor as Virtual Archaeologist: Explorations in Mixed Reality Technology to Enhance Educational and Social Interaction in the Museum, ACM

conference on Virtual reality, archeology, and cultural heritage.

Kannan, R., F. Andres and C. Guetl, 2010. DanVideo: an MPEG-7 authoring and retrieval system for dance videos (pp.545–572) Multimedia tools and Applications.

Kaufmann, H., and B. Meyer, 2008. Simulating Educational Physical Experiments in Augmented Reality, ACM SIGGRAPH ASIA educators programme.

Kaufmann, H. and D. Schmalstieg, 2002. Mathematics and geometry education with collaborative Augmented reality, ACM SIGGRAPH conference.

Kuramoto, I., Y. Nishimura, K. Yamamoto, Y. Shibuya and Y. Tsujino, 2013. Visualizing Velocity and Acceleration on Augmented Practice Mirror Self-Learning Support System of Physical Motion, International Conference on Advanced Applied Informatics.

Majumdar, R., and P. Dinesan, 2012. Framework for teaching Bharatanatyam through Digital Medium (pp. 241-242), IEEE Fourth International Conference on Technology for Education.

Randy, W.H., B.M. Pauline and K.T. Danesh, 2001. Visbench: A framework for remote data visualization and analysis (pp.718–727), Proceedings of the International Conference on Computational Science-Part II, ICCS '01, London, UK Springer-Verlag.

Silva, P.A., F. Nunes, A. Vasconcelos, M. Kerwin, R. Moutinho and P. Teixeira, 2013. Using the Smartphone Accelerometer to Monitor Fall Risk while Playing a Game: The Design and Usability Evaluation of Dance! Don't Fall (pp. 754–763) Foundations of Augmented Cognition.

Yang, Y., H. Leung, L. Yue and L. Deng, 2012. Automatic Dance Lesson Generation, IEEE Transactions on Learning Technologies.