2018 8th IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2018), 23–25 November 2018, Penang, Malaysia

Exploring Dancer Experience and Performance with Visual Sensor Camera Using NUI Environment

Manjit Singh Sidhu
Dept. of Graphics and Multimedia
Universiti Tenaga Nasional
Kajang, Selangor, Malaysia
manjit@uniten.edu.my

Javid Iqbal
Dept. of Graphics and Multimedia
Universiti Tenaga Nasional
Kajang, Selangor, Malaysia
manjavi@gmail.com

***Abstract*—Visual reality sensors with high resolution depth (RGB) can make the computer see things and provide interactive feedback like humans therefore providing opportunities to resolve elementary issues in computer vision (CV). In this paper we discussed a novel system that detects dance motion and analyzes the dancer's dancing steps using natural user interface (NUI) environment. Pre-recorded videos of a particular activity such as dance which are stored in a database are used to provide training to the users that are then compared and analyzed by the system providing feedback of the postures in real-time. The results of our pilot test on the design of the virtual analysis system that is capable of improving, providing timely feedback and reveals a strong motivation from users wanting to learn and enhance their dancing steps and postures of the particular selected dance.**

***Index Terms*—Analysis, Feedback, Interactive, Motion, Visual reality Sensors.**

I. Introduction

While the last decade was dominated by interactive technologies, the next several decades will be dominated by visual reality sensors. Technological advancements such as visual reality sensors have enabled computer systems to process tasks similar to humans. Thus it is possible to make the computer see things and provide interactive feedback like humans hence shaping the new reality and providing opportunities to resolve elementary issues in CV. For example visual reality sensors can be used to detect human postures and motions to train humans for fitness training, learning how to perform yoga and martial arts or even practice a sports event without having to attend training at specialized training centers. Fraser stressed that in many physical activities, mastering motions and new postures are crucial components [1]. Some categorize this sort of work as augmenting human [2]. Natural user interfaces (NUI) such as speech and gestures are also becoming useful as a means of communication with these systems. Since the integration of CV with NUI for PCs has not been tapped for its effectiveness, our present research focused on the design and development of a system that records and analyses the dancing steps and postures with this approach.

Dance is better known as human body motion or gestural communication with music [3 – 5]. However an essential part of dance is to form a charming movement of dance flows to generate a visually engaging sequence of actions [6].

Kar [6] stated that dance is explicit to diverse regions, communities and personal styles. Dance institutions around the world teach totally different dance moves to students, whereby they are typically needed to mix to create new forms of dance sequences. It is a vital a part of any dance creation for trainers and self-learning for learners.

Although numerous techniques have been planned to ease dance choreography, most of those works are tested on the 'Ballet' dance [7] & [8]. As such we experimented more varieties of dance styles namely the Indian "Bollywood", "Baratnatyam", "Western", and a popular Malaysian folk dance "Joget Zapin". We employed a newer form of blended technology using computer vision and augmented reality (AR) to create a sense of mixed reality of a new interactive experience that are human centred. AR has been defined by technology pionner Azuma (1997) as "*AR allows the user to see the real world with virtual objects superimposed or composited with the real world*". According to Danakorn, new possibilities for enlightening the quality of teaching and learning activity has made it possible with the ability of combining virtual and real worlds together to create new prospects [9].

Conventionally the AR technology works by tracing a target in the real world by using a camera and software on an enabled device such as the PC or smartphone. These targets can include things like icon, an image, an object, a sound, a location or even a human. The target input data is processed by the software and compared against a database of potentially corresponding information. If a match is identified, an AR experience is triggered and content is superimposed on top of reality.

II. Background

Before the existence of personal computer technology, trainees may learn dance by watching educational CDs/ DVDs. However the later trend has been overcome by the supplementary of online video web sites, like YouTube that provides coaching videos with thousands of movements, for

978-1-5386-6324-0/18/$31.00 ©2018 IEEE

varied skill activities including sports. Although videos for learning the skills are available on the market, the utilization of the videos as a coaching medium is restricted [1]. Videos limit feedback, unable to capture 3D movement information and have restricted interactivity and do not provide customized motivation. There has conjointly been associate emergence of fitness and dance video games that use game consoles with tracking technology [10]. Although existing games might improve some motion ability, their target has centered on amusement and physical exercise therefore movements are strained to pre-programmed sequences, and do not offer smart feedback.

There has been extensive research in the area of human detection from still and video images. Common usage includes histograms [12], Partial Least Square (PLS), which are all based on colour images from visible light [13]. Kuromoto [14] anticipated an image technique of speed and acceleration of instructor's motion for the user to comprehend straightforwardly. However, with the arrival of AR technology with visual reality sensors, many industries have diverted their attention around the globe creating useful applications for problem solving. In the health industry, medical practitioners can interact with virtual 3-D models of human anatomy. Additionally 3-D print of different procedure stages can be manipulated. Papagiannis predicted that new developments in haptics will enable a surgeon to day work on virtual brain physically, engaging in a full tactile experience before performing an actual surgery [2]. AR also offers instructional repair guidance in the manufacturing industry. It would be possible to see, share and receive information in real time collaboration and explanation. For example an architect in Korea could be on location with a builder in Malaysia, interacting and fully engaged at the job site [15].

In the entertainment industry, a television may not be required anymore, the AR headset could be used as the entertainment hub allowing innovative types of digital content to be tailored to and coincide with the physical surroundings. Recent advancements in applications development have blended AR with computer vision to another level. For example a particular task could be done by a human both locally or remotely and the computer is made to see with a visual reality camera, analyse and instantly provide useful feedback [16]. In the physical environment, some systems were designed using Microsoft Kinect that can be used with the Xbox and TV for example, "Dance Central 3 Xbox 360 Game for Kinect" as shown in Figure 1. In this system users can learn the popular dance from across the ages, including disco moves of the 70s and hip-hop moves from the 80s.

Fig. 1. Classic dance moves from past decades.

Another system "Insanity" which include authentic dance moves that are fun for dancers of all skill levels, targeted for working up for sweat, perfecting a performance and parties. As for fitness systems "Work it out" is one example that gets the heart pumping with expanded and enhanced calorie tracking features and premade fitness playlists that make it to fit in a quick workout while dancing to the desired songs. Figure 2 shows an example of a fitness trainer in training people how to lift weights or exercise.

Fig. 2. Fitness training to lift weights and exercise.

As mentioned earlier in this paper, we focused our attention of four selected dance forms to see the desirability of using a virtual trainer with natural user interface interaction (NUI) for tracking human motion expression which could automatically be analyzed by a computer system. Further on the design and development is discussed in the next section.

III. System Design

Since computers can be used to create software that exhibits computational creativities [17], we developed an interactive dance analysis system to help enthusiastic dance learners to enhance their dancing skills. The conceptual system architecture is shown in Figure 3. According to Figure 3, the system enables a student to learn the selected dance style by viewing the virtual choreographer (VC) on the projector screen where the VC performs the dancing steps based on the dance style selected by the student from the motion database. When the student follows the instruction of the VC, the dance motion is detected and captured by the motion detector and analyzed by the emotion analyzer which then provides interactive feedback to the student. In order to make the student get a real feel while dancing, AR is used to include realistic background, visual effects, acoustics, dancing of an expert choreographer according to the dance style that will hence create motivation. The emotions, dancing steps and postures are analyzed by the visual reality sensor.

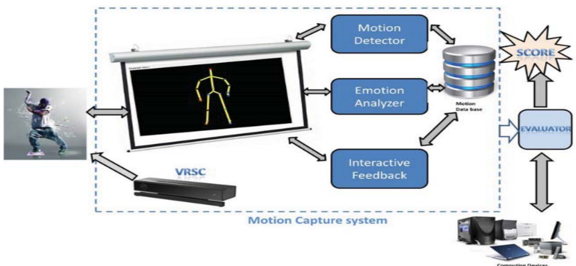

Fig. 3. The system architecture.

A. Metric for comparing the dance videos

The fundamental concept in placing dance-videos together in a visually appealing dance choreography relies on a straightforward comparison metric of these videos. The dance video clips of variable duration ranges from 2 to 5 minutes (basic - advanced). Visual reality camera sensor devices normally records at the speed of about thirty frames per second, therefore the amount of frames captured for each dance-video is different. A number of frames in each video clip is decoded into a sequence of skeletal diagrams, where each skeletal diagram corresponds to a particular frame (illustrated in Figure. 4). The skeletal diagrams depicted in Figure 4 are sketches of the body structure shown by 3-D straight line segments connecting 20 essential intersections of the dancer's physique (Figure. 5). Each 3-D straight line segment is projected on to XY, YZ and ZX planes as shown in Figure. 6 a, b, c, and the slopes of the projected 2-D straight lines in the three planes with respect to X-, Y- and Z- axes correspondingly are calculated. These angles are used as metrics to relate two different 3-D straight lines representing the orientation of one given body part (i.e., the right forearm) present in two frames of two videos. Logically, the matching of 3-D straight lines is needed between the last frame and the first frame of each pair of videos to test their possible just-apposition within the final video.

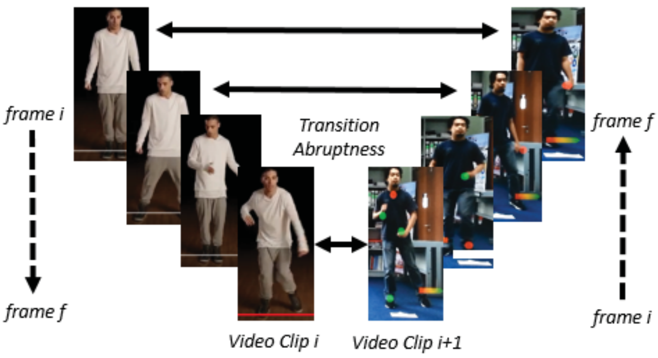

Fig. 4. Fig. 4. Dance videos (video clips) demonstrating frames. Transition Rapidity is computed by matching the end and beginning frames of consecutive dance videos i and i + 1 correspondingly.

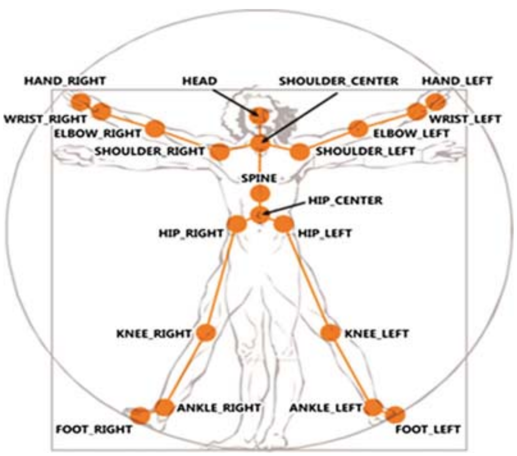

Fig. 5. Body join coordinates from visual reality sensor camera for each frame.

TABLE I
PERCENTAGE FREQUENCIES OF STUDENT'S PREFERENCES BETWEEN THE VIRTUAL AND REAL CHOREOGRAPHER

Choreographer	Virtual	Real
Preferences	58.1	41.9
More effective	66**	34
Easier-to-follow	60.1*	39.9

*** p<0.05; ** p<0.01**

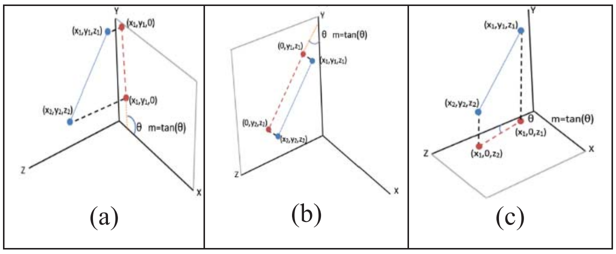

Fig. 6. Computation of slope of projections of a line on XY, YZ and ZX planes X, Y and Z axes correspondingly.

In terms of comparing and analyzing the dance steps, a metric for comparing the dance steps of the particular pattern i.e. western and to use it in comparing with the dance motion by the dancer (student/trainee) is proposed. Here depending on the level selected by the dancer (basic – advance), each dance step labeled with a marker point mapped to the virtual choreographer in the video is compared with the dancer in the real time environment. While the algorithm computes and processes the steps, the system analyzes and provides immediate feedback by acoustics which are also supplemented with visual cues for motivational or improvement.

B. System testing

The system testing of this research was done at the advance technology assisted problem solving lab (ATAPS) at, Universiti Tenaga Nasional (UNITEN). A visual reality sensor camera (VRSC) was used which can capture skeletal co-ordinates at an estimated rate of thirty frames per second. Therefore in closed room settings, this type of configuration allows the VRSC to efficiently capture the 3-D skeletal co-ordinates.

Certain precautions are also needed to ensure noise free collection of data. For example the dance trainees are told to avoid wearing clothes which are too loose around the body joints. The dance trainee should be within an estimated range of the VRSC i.e. approximately 1.2 - 3.5 meters or 3.9 - 11 ft. Additionally the dancers should avoid wearing black colored clothes, shoes and socks to ensure proper mapping of points from the video to the dancer so that the mapping points are appropriately initiated in the real-time environment.

When the visual cue indicating "*please follow*" is shown on the screen, the dancer should then follow the dancing steps of the virtual instructor shown on the top left video in the window. The red points shown in the figure are used to map the skeletal joints with the virtual choreographer. The pointer labeled with three colors shown in Figure 7 (mapped to the foot) shows the next step to be done by the student which further indicates the perfection with green being correct and red being wrong and yellow somewhat close while scores are shown on the right side of the screen indicating how well is the dancer performing. The video clips of the dynamic dance motions were also being automatically recorded and stored in a database which could be visualized by the student to see how they had dance.

Regular testing of the interactive analysis system was conducted. Major testing aspects of the system included; recording and ensuring the correct dance steps of the dance experts, mapping the dance steps with the skeletal body join coordinates of the visual reality sensor camera for each frame and the algorithm for providing feedback in real-time environment.

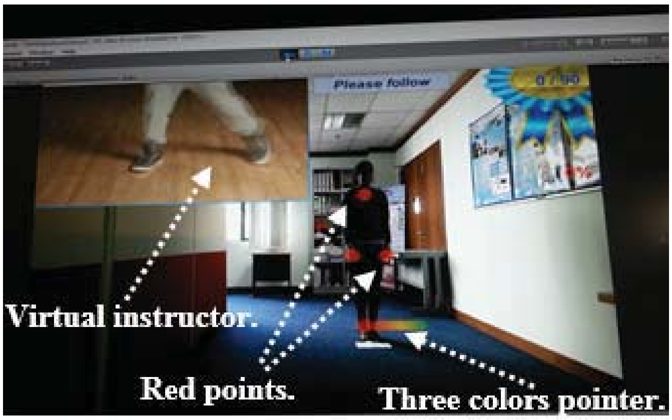

Fig. 7. Student performing western dance.

IV. Methods, Results And Discussion

The subsequent sections describe the approach taken in evaluating the interactive dance analysis system.

A. Participants

The study was conducted during the second semester period (2017). In our initial pilot system test, 10 students were asked to perform the western dance which was captured separately each one with the help of the VRSC (Figure 8). They were also asked to fill in a self-designed, five point Likert-point questionnaire. The aim of the pilot test was to see if the skeletal points were mapped correctly; students could follow and understand the visual cues; the suitability of the virtual instructor video window size; and the duration (time) to learn the particular dance. The questionnaires were tested for its suitability of contents and time to answer the questions. Although the results did not permit meaningful quantitative statistical results, the findings are of an empirical experiential nature and we found the expressions of the participants were overwhelmingly consistent.

Fig. 8. Arrow at foot step indicating perfection.

B. Setting

Practical experiments were equipped with a laptop and the VRSC. The subjects (students) were first briefed about the project and their consent was obtained. The students were then asked to stand at the appropriate position with the correct posture and follow the dancing steps of the virtual instructor. The time duration depended on the option selected by the student i.e. basic – 4 minutes, intermediate – 3 minutes and advance – 2 minutes). After completing the dance, each student was asked to complete the 5 Likert-point questionnaire comprising of 50 questionnaires. All the questionnaires were completed and returned to the researchers. The questionnaires were statistically analyzed using SPSS ver. 22.

C. Design

In this study, we used quantitative data by means of statistical tests by designing a two-way ANOVA. Students' preferences (SP), effectiveness and easiness-to-follow are the dependent measures of the study. Additionally the study consist of 3 distinct phases namely (1) recording of the participants' profile, (2) interaction with natural user interface and visual reality sensor feedback, (3) free sequential interaction with dance analysis system and completion of the questionnaires. Apart from that, data from observations and interviews were analyzed.

The practical lasted about 2 hours for all the 10 students (this includes explanation by the researchers on how to use the dance analysis system and filling up the questionnaire). The students were also asked about their preference in learning dance from a real or virtual choreographer. The answer to this question was recorded as the preference towards the preferred method.

V. Qualitative Data Collection Method

A dual methodology approach was employed to collect qualitative data. The first one is structured observation and the second one is student interviews. All data on the other hand were collected by the researchers in the form of observation notes. In more detail:

(i) **Structured observation:** Two researchers collected data that emerged from student's interaction while they were interacting with the dance analysis system. The data collected were from: (1) the discussions between the students, (2) personal observations that focused on the difficulties that the students confronted.

(ii) **Interviewing:** upon completing the questionnaire, students were interviewed based on their knowledge (semi structured interview). This approach was adopted to further confirm and understand the data collected from the questionnaire and the observations. As such, students were requested to provide more details about their knowledge and comment on the points that needed explanation. The data that was correlated with the previously captured data provided opportunity for confirming and fine-tuning the resulting conclusions.

VI. Qualitative Measurements

C. Results – Quantitative analysis

The percentage frequencies of the student's preferences regarding the virtual and real choreographer's feedback of the dance analysis system is shown in Table 1. Based on the sample, the virtual feedback was more prone to be favoured as the student's preference (however without statistical significance), while it was characterized as more effective and easier to follow (both statistically significant, $p<0.01$ and $p<0.05$, respectively).

D. Preference/Effectiveness Variable

Students showed statistically significant higher preference for the virtual choreography method ($x^2 = 4.12$, $p<0.05$). The effectiveness was measured at two successive points in time: (a) before direct interaction, and (b) at the end of all hands on practical session. The findings revealed that preferences was higher during the practical training for the virtual choreographer and these differences are statistically significant with chi-square ($p<0.05$) in all cases except effectiveness for the students who favored the real choreographer.

E. Ease-to-follow Variable

For the ease-to-follow-variable the virtual choreographer method was found to be easier to follow for the students and which provides evidence that this effect is changing with newer technological enhancements. The results revealed the effect, which was statistically significant, (Cramer's V = 0.62, $p<0.0001$). Since these effects on virtual choreographer were measured comparatively to the real choreographer, any +ve or -ve changes throughout the groups of the independent variables could be considered as -ve or +ve alterations, respectively, for the virtual choreographer.

F. Observation and Interviews

From our observation and interviews, all students generally agreed that both methods (real and virtual choreographers) worked very well and helped students to successfully learn, interact, improved their dance steps of the selected dance style. In order to ensure that the students were satisfied by the outcome, the assigned tasks were on purpose selected to be easily completed by the them within the given time frame. The objective to complete successfully the dancing steps was met for all the students and conditions. On the other hand most of the students appear to be attracted to use the new dance analysis system with the visual reality sensor and were engaged with the dance learning steps. In fact the students showed high interest particularly during the free testing period to learn all the dance styles offered and features of the system. The allocated tasks were on purpose selected to be easily achieved by the students, within the available time, so that they were successful and satisfied by the outcome. Indeed, the goal to complete successfully the dancing steps was fulfilled for all the students and circumstances. Generally, most of the students seem to be attracted to see the new dance analysis system with visual reality sensor and were engaged with the dance learning steps. The students indicated high interest to learn all the available dance styles and features of the system, especially during the free testing period.

G. Discussion

In this paper, we present a dance analysis system application that uses visual reality sensor to analyze students dancing steps to see if computer vision could be used as an effective training tool. The results showed how the dependent variables might be varying by student's preferences of the two preferred dance choreographers namely virtual or real, when they interacted using the natural user interface. The aim of the research is to make a contribution to the ongoing investigations on AR based applications and encourage researchers to further explore specific contexts in computer vision in training particularly where human expertise is limited. Additonally to expose the new technology to users.

H. First-sight-preferences

The results of the quantitative data supported by the observation and interview, shows that the virtual choreographer is more preferred by the students. The quantitative results of this study are also consistent with another study, virtual choreographers were preferred [6].

I. Effectiveness

Regarding effectiveness, student's self-reports indicate that the dance analysis system is effective. This evidence was equally strong and it is reliable in all measurements that took place during the testing process. Particularly notable is the application whereby students mentioned that they could concentrate on the virtual choreographer and it was similar to a real choreographing concept. Additionally commented that they were not afraid to make mistakes in the dancing steps. Among other observations were that students could easily use the NUI to select other dance if they wished and they could learn dance at their own pace.

J. Easy-to-follow

Regarding easiness-to-follow, the results appealingly showed that in general the virtual choreographer is assessed as accepted because constant feedback is given during the dance and it was motivational.

VII. Conclusions

In this research, we were looking at novel ways to present dance training for beginners and create a motivational environment to enhance their dancing steps without having to attend traditional dance classes. Testing the effectiveness of a blended AR technology with visual reality sensor camera training system seems like a natural step in evaluating the benefits of such systems. Therefore it can be noted that there are features available in such systems that are not available in present training methods, so there is bound to be utility in CV for training.

This study explored student's opinions and preferences regarding a dance analysis system with visual reality sensor. The results indicated that the virtual choreographer used in the system was acceptable especially for teaching the dance steps, was effective and easy to be follow as the students could easily interact with the system. Factors that appear to have influenced student's preferences embrace the preceding familiarity with the popular Xbox game ideas. It is viable to perform the dance with a virtual choreographer and produce the results in the real environment and combine the dance steps with the real ones.

The pilot test showed that the blended AR technology with visual reality sensor camera has the potential to train students from a virtual instructor that provides timely feedback that mimics the conventional way. The test showed that the interactive dance system was intuitive and easy to use and navigate as the students did not approach the developers much for help in interacting with the system. In order to foster learning in real time dance environments, and improve the visual view, large screen would be beneficial as the space on small screen may hinder the students from seeing the videos clearly. The overall results could be concluded that most of the students agreed with the usefulness, effectiveness and ease-to-follow of the dance analysis system.

We predict that it may not be long when we see virtual instructors fully taking over the job of a choreographer in teaching how to dance, a fitness trainer in training people how to lift weights or exercise correctly and a yoga teacher in teaching the correct steps in performing to improve body flexibility, balance and strength.

It is assumed that the system will be a valuable coaching software tool for many circumstances, just as virtual environments have shown to be successful. Through a systematic process of testing what qualities are beneficial and what features obstruct the training, we look forward to bring AR and visual reality sensing systems to a level which will allow other sorts of training where real human experts are limited to learn to integrate them into effective training tools. We have perceived some factors for future studies in executing the experiments we have stated. For instance the reliability or realism in the static and dynamic models, lighting to match the real environment, and occlusion by real objects. The level of reliability needed for quality training is the aim for future work.

Acknowledgment

Thanks is due to Universiti Tenaga Nasional (UNITEN) for supporting us with the Uniten Internal Research Grant Funding No: J510050729 (UNIIG 2017) to conduct this research.

References

[1] A. Fraser, T. Grossman, J. Matejka, & J. Fitzmaurice, "YouMove: Enhancing Movement Training with an Augmented Reality Mirror Dance learning and the mirror: comparison study of dance phrase learning with and without mirrors." *Dance Education,* UIST'13, ACM, 2013 pp. 311-320.

[2] H. Papagiannis, "The book, Augmented Human, How Technology is shaping the New Reality." O'Reilly ISBN: 978-1-491-92832-5. 2017.

[3] B. G. Patra, D. Das, & S, Bandyopadhyay, "Unsupervised approach to Hindi music mood classification." *MIKE. LNCS*, Springer, Heidelberg.vol. 8284, 2013, pp. 62–69.

[4] I. Javid, & S. Manjit, "A System Framework for Computer Vision and AR based Dance Learning Technology." *Aust. J. Basic & Appl. Sci.* 9(19): 2015, pp. 27 – 34.

[5] R. Rimmer, "Negotiating the rules of engagement: exploring perceptions of dance technique learning through Bourdieu's concept of 'doxa'." *Research in Dance Education* 17(3) 2017, pp. 221 – 236.

[6] R. Kar, A. Konar, & A. Chakraborty, "Dance Composition Using Microsoft Kinect." *In: Trans. of Comput. Sci*, Springer International Publishing. 2015, pp. 20–34.

[7] A. Soga, B. Umino, M. Hirayama, "Automatic composition for contemporary dance using 3D motion clips: experiment on dance training and system evaluation. In: International Conference on CyberWorlds." 2009 IEEE. pp. 171–176.

[8] J. Dancs, R. Sivalingam, G. Somasundaram, V. Morellas, N. Papanikolopoulos, "Recognition of ballet micro-movements for use in choreography. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)." 2013. IEEE. pp. 1162–1167.

[9] N. Danakorn, B. Mohamad, H. Noor, R. Hishamuddin, "Mobile Augmented Reality: the potential for education International Educational Technology Conference." 2013.

[10] Dance Central. 2010. Harmonix, Massachusetts, USA.

[11] M. Sims, A. Mark, J. Clasey, S. Beighle, A. Fedawa, & H. Erwin, "Development of a system for observing dance activities in the classroom environment (SODANCE). Research in Dance Education" 17(3) 2016. pp. 161 – 175.

[12] N. Dalal, B., Triggs, C. Schmid, "Human detection using oriented histograms of flow and appearance, European Conference on Computer Vision, Graz, Austria." 2006.

[13] W. Schwartz, A. Kembhavi, L. Davis, "Human detection using partial least square analysis." ICCV. 2009.

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

[15] Goldman. Equity Research. Goldman Sachs Global Investment Research Report. 2016.

[16] Nabshow. "Where content come to life. Trends in Broadcast." 2017. pp. 1- 22.

[17] S. Colton, R. López de Mantaras, O. Stock, "Computational creativity: Coming of age." AI Mag. 2009. 30(3), 11–14.