2022-08-31

DanceShala - A Visual Feedback Interface for Dance Learning

Mukherjee, Suvojit

UNIVERSITY OF CALGARY

DanceShala - A Visual Feedback Interface for Dance Learning

by

Suvojit Mukherjee

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN COMPUTATIONAL MEDIA DESIGN

CALGARY, ALBERTA

AUGUST, 2022

Abstract

Dance is a beautiful art form that can be enjoyed by people irrespective of age. One can learn dance from a dance teacher in a dance studio. The visual feedback received in an in-person class from an instructor is one of the best ways to improve dance learning. Sometimes it is not possible for a person to attend dance classes due to time and location constraints. The alternate option for people to learn dance is to attend dance classes online or self-learning with the help of dance games (online video games where a player attempts to follow a pattern of dance steps shown on screen in time to music). Organized remote visual feedback can assist a learner to learn dance in such scenarios. However, online dance classes or dance games may not be sufficient for a new learner to learn dance because the feedback received is not always adequate. To make online dance learning more comprehensive for new learners, a visual feedback interface named 'DanceShala' is created which will deliver comparative visual feedback to students after comparing teacher and student movements. In this study, dance movements of the teacher and the student are recorded. After processing the recorded movement data, feedback is generated on the correctness of the student's movements as compared to the teacher. The visual feedback is displayed through an interface which assists a student to identify the errors made when compared to the teacher video. In the last stage of the study, a survey is administered to understand the user perception about this interface. This research is an interdisciplinary study combining Computer Science, Kinesiology and Dance.

Acknowledgements

First of all, I would like to thank the University of Calgary's Computational Media Design (CMD) department authorities for giving me the chance to be a part of the program. The CMD department brings together a small, diverse group of students and staff members that accept out-of-the-box ideas and are not afraid to push outside their comfort zones. It is quite difficult to gather scholars from different disciplines under the same umbrella unless the CMD authorities have a crystal-clear approach. For that reason, I'd like to express my gratitude to the CMD team, particularly the recent director, Dr. Lora Oehlberg, as well as CMD's former directors, Dr. Patrick Finn and Dr. Jeffrey Boyd, for their long-term dedication and significant effect on the program. It was a true honor for me to be a part of this program.

Progress in a cross-disciplinary Masters, is always a challenge which would have never been possible without the support of my supervisors. On that end, I would like to thank my supervisors Dr. Usman Alim from Computer Science department and Dr. Sarah Kenny from the Divison of Dance for assisting me throughout the program. Dr. Usman Alim is a gem of a supervisor one can ever get. He has been helping me since the start to direct my thesis to the proper direction. Dr Usman's solid supervision really enriched not only my thesis, but also my academic development. Dr. Sarah Kenny has always been supporting me with her ideas and views which has really helped me to bridge the gap between dance and computer science. Her opinions and valuable feedback about the artform have helped to shape my thesis in a proper constructive manner.

I am grateful to the entire Visualization and Graphics Group, aka VISAGG team. I'd like to thank Allan, Mamun, Roberta, Julio, Mohammad, Mahnaz, and Amin for their friendly, patient, and highly helpful comments on my study. I extend my sincere thanks to Haysn and Joshua for their extensive help and support throughout my entire masters.

I was most honored to have Dr. Aditya Shekhar Nittala and Dr. Larry Katz as my examiners. I want to thank these wonderful people and scientists not only for their time and presence at my defence session, but also for their valuable ideas, and insightful outlook on my work.

My final round of appreciation goes to the all the participants who have who helped me with pilot testing by participating in the study. I am very grateful to all of you.

Table of Contents

List of Tables

List of Figures and Illustrations

List of Symbols, Abbreviations and Nomenclature

Symbol	Definition
U of C	University of Calgary
HCI	Human Computer Interaction
VR	Virtual Reality
UI	User Interface
BVH	Biovision Hierarchy Data file

Chapter 1

Introduction

Historically, dance has played a role in social ceremonies, spiritual practices, social entertainment, and healing rituals [1]. Art was thought to cure both the individual and the group in old societies. It is clear that dance embodies all facets of humanness, from ancient artefacts to those of current society. Dance may now be seen by the whole public as an art form, a means of physical conditioning, a type of social entertainment, a therapeutic aid, or a contribution to educating creative citizenship in public schools[1].

In this research, dance and technology are combined to develop 'DanceShala', a visual feedback interface which assists a beginner to learn how to dance in an online format. In Hindi language, 'Pathshala' means school. As this is a dance learning platform, it is named 'DanceShala'.

In this introductory chapter, motivation and purpose in developing this project, thesis statement and thesis contributions are discussed. In addition, a summary of the application's design and engineering is also presented.

1.1 Motivation

Dance is a fundamental art form, involving expression of the human body using movement [2]. Dance helps to improve both physical and mental health [3]. This research investigates on how to make dance learning easier for beginners.

In 2022, a person can learn dance online as well as through in-person classes. Offline or in-person dance classes are a great way to learn dance. Sometimes due to time and location constraints, people are not able to avail the opportunity of in-person classes. In these cases, people resort to online options of learning dance where technology plays a major role. In the world of dance education, recent advancements in technology have resulted in the 'design and development of tools to support dance teaching and learning through different technologies such as interactive whole-body experiences, motion capture, as well as virtual and augmented reality' [4]. Usui et al. [5] suggest that one reason to encourage online interactive dance learning is that young teachers can utilize technologies as supporting tools to get rid of teaching anxiety and lack of teaching experience. Motion capture and animation technologies could potentially lead to the development of new and exciting methods to help people learn dance. The usage of interactive technologies to support dance learning and education is still a relatively under-explored domain [6]. The technology utilized until now in dance has certain limitations. Sometimes they are very complex, and sometimes they are too expensive to afford.

Another aspect that is important in dance learning is feedback. A skill can be improved upon by feedback. Dance games like Dance Dance Revolution [7] and Just Dance [8] may help to learn dance to some extent but the feedback provided by these dance games is limited likely because the game is more focused on the entertainment and 'fun' factor. More specifically, in these cases, Chan et al. [9] indicates that the input data is greatly reduced to simplify the movement analysis and although such decimation is justifiable in an entertainment setting, the system is not able to dispense sufficient feedback for the whole-body movement for training purposes. Also, according to the results presented by Charbonnau et al. [10], dance games do not seem to offer much in improving dance performance and learning, as usually the feedback is just a judging overall score, which neither indicates what went wrong nor offers strategies for improvement. There are also many communication services like Zoom [11] and Google Hangout [12]. As these platforms are mainly designed for meetings, there can be technological challenges in viewing and analyzing movements due to a static camera position. In addition to that, it becomes harder to concentrate on movements of a single person in a group class in Zoom [11] or Google Hangout [12] as a teacher needs to concentrate on multiple windows at a time.

1.2 Goals

There are numerous challenges like portability, complexity and user-interactivity in the currently used methodologies utilized for dance learning. These shortcomings informed the actual goals for this research which is to provide assistance to beginners in dance learning by utilizing less complex technologies. Keeping this goal in mind, a simple interactive dance visual feedback interface is created that strives to make dance learning accessible for anyone interested to learn dance.

1.3 Procedure

The methodology employed in this work consists of three steps: 1) capture a beginner's dance movement, 2) compare it with a teacher's similar dance movement, and 3) present feedback to the user in a visually appealing and understandable manner. The motion capture is done with the help of simple equipment (i.e. a mobile camera). After motion capture, motion data is extracted from the videos using Radical AI [13]. Then, motion analysis is performed on extracted data to detect movement dissimilarities comparing bone positions of the teacher and the student. Afterwards, feedback is presented through the visual feedback interface. Finally, a survey is conducted with experienced dancers for feedback on the interface and identify future enhancements.

1.4 Scope and Contributions

The main contribution of this thesis is 'DanceShala', an application whose primary purpose is to contribute to the field of dance education by combining dance and computer science in a novel manner. Specifically, the contribution of 'DanceShala' in the field of dance education as an affordable, portable, and simple software is worth mentioning. 'DanceShala' has been specifically designed for beginners who have a passion to learn dance, but sometimes are unable to attend in-person dance sessions owing to a variety of factors. 'DanceShala' can also be utilized by dance teachers to deliver visual feedback to beginner students.

1.5 Thesis Organization

This thesis explains the methodologies and techniques that have been used to design, construct, and evaluate 'DanceShala'.

Chapter 1 (Introduction), as already discussed, contains the motivation and purpose for developing this project. This chapter also highlights the scope and contributions of the research. Apart from the introduction, this document is subdivided into five chapters.

Chapter 2 (Background) provides an overview of the various kinds of dance feedback, motion capture methodology and equipment, and various dance genres. This chapter contains a review of relevant literature to provide context for the research topic.

Chapter 3 (Design Choices) explains design choices for this research.

Chapter 4 (Setup) describes the stages involved in motion capture and motion analysis, software and hardware employed, and an explanation of the functionalities implemented.

Chapter 5 (Visualization Interface) explains in detail how the interface was created and what the important features are.

Chapter 6 (Survey) contains information related to the user survey which was conducted to ascertain the efficacy of 'DanceShala'.

Chapter 7 (Conclusion and Future Work) provides an overview and assessment of the work completed, acknowledgement of the project's limitations and potential future research possibilities.

Chapter 2

Background

This section provides background details to help readers understand some of the study's key themes. Firstly, this chapter covers online and in-person dance learning, different types of feedback, effectiveness of feedback, and feedback in online and offline dance instruction. Secondly, it discusses classic motion capture and motion matching approaches, as well as their obstacles and limitations. Finally, an overview of various dance forms is provided.

2.1 Online and In-Person Learning of Dance

One of the reasons for developing and deploying interactive dance learning systems is to allow for remote learning. But, when it comes to dance instruction, is online learning a feasible option? Online dance learning shares the same characteristics with online physical education. For instance, Daum et al. [14] state that dance and physical education face similar challenges as other topics in online learning courses. Academic integrity, learner preparedness, motivation, student retention, and technological challenges are just a few of them. Furthermore, as described by Camurri, in the framework of WhoLoDancE (Whole-body Interaction Learning for Dance Education) [15], these movement-oriented themes present the learner with the unique challenge of improving motor skills, movement principles, and attributes in a virtual format.

Buschner and Craig [16] outline some of the possible benefits and drawbacks of online physical education, which might be applied to online dance instruction as well. The key advantages are that younger students may be driven by technology as a novelty in some situations. Most importantly, this online dance education can help those students who reside in rural places without easy access to dance schools to learn dancing.

Recently, COVID-19 pandemic has changed the way of learning dance [17]. Teaching and learning activities that are usually carried out with in-person had to adapt to virtual online applications [17]. As such, striving to enhance the online dance learning experience is needed [18].

2.2 Feedback

Feedback is defined as information on one's performance or knowledge offered by an agent (e.g., a teacher, a peer, a book, a parent, oneself, or experience) [19]. In this study, though the feedback is interpreted as a result from a physical movement performance and doesn't encompass the artistic expressions entirely, still broader feedback on the whole dance performance is provided.

2.2.1 Importance of Feedback

Feedback is acknowledged as an important component in improving students' learning processes [20]. Feedback is widely regarded as an important tool for assisting students in their growth as autonomous learners who can monitor, analyze, and manage their own learning [21]. It is recommended that supportive feedback include the following factors:

a) Student's Learning Goals

Learning goals describe an intended purpose or desired outcome that specify the information, skills, and capabilities that the student is striving for. If a student can understand or decide on their learning goals, it will be easier for them to attain those goals [22].The students must experience a sense of ownership over the learning goals and be able to understand self-evaluation [21]. Research demonstrates that there are several gaps between teachers' and students' perception of goals [23], evaluation criteria, and standards. The goals specified by students and the goals initially stated by the teacher should have a fair degree of resemblance.

b) Simplifying the Improvement Process of Self-Evaluation in learning

There has been an increasing interest in self-evaluation during the last several years [24]. A technique to strengthen self-evaluation is to provide the students the opportunity to practice and reflect on controlling qualities of their own learning. Students are frequently involved in observing discrepancies between internally defined task goals and the outputs that they produce [23]. Feedback can be effective if self-evaluation is developed.

c) Providing Quality Information to Students about their Learning

Teachers play a critical role in enhancing their students' comprehension of the self-evaluation process [22]. They're also an important source of outside input. Feedback from teachers has traditionally served as a means for students to assess their development. The students can also monitor their own personal improvement by the feedback from teachers. Furthermore, teachers are frequently more effective than students at recognizing errors in their work [22]. As a result, it is critical to provide students with high-quality material in order to ensure that they learn [22]

d) Allowing Teacher-Student Dialogue in Understanding the Feedback

One way to improve the value and efficacy of feedback, as well as the chance that students will understand the information presented, is to think of feedback as a dialogue rather than a transmission of information [23]. The phrase 'feedback as conversation' refers to the fact that the student will not only get written evaluation, but will also have the option to discuss it later [22]. In this case, in order for feedback to be effective and helpful, it must first be comprehended by the student before it can be used to create beneficial changes.

e) Motivation

Motivation is a complicated aspect of human psychology and behaviour that impacts how people choose to spend their time, how much energy they put into an activity, how they think and feel about it, and how long they stick with it [25]. A students' motivation is reflected in their choices of learning activities, the time and effort they dedicate to them, their perseverance on learning tasks, and how they deal with hurdles they face during the learning process [25]. In both learning and evaluation, motivation is critical. A teacher should understand the students' needs, interests, purposes, and attitudes so that they are able to encourage the students to learn. Learning may not occur if students are not properly motivated [26].

f) Providing Opportunities to Close the Gap between Current and Desired Performance

Feedback may help the students to bridge the gap between their intended and present performance. Boud [27] is of the opinion that the only way to tell if learning results from feedback is for students to make some kind of response after receiving feedback [28]. Unless students are able to use the feedback to produce improved work, neither they nor those giving the feedback will know that it has been effective.

g) Providing Effective Information to Teachers which can help to Shape the Teaching Practice:

A good feedback practice may not only provide helpful information to students to help them enhance their learning, but it can also provide useful information to teachers, resulting in a better learning experience for students. Yorke and Mantz [29] have the following opinion regarding this issue: 'The act of assessing has an effect on the assessor as well as the student. Assessors learn about the extent to which they [students] have developed expertise and can tailor their teaching accordingly.'

2.2.2 Feedback in Dance Learning or Dance Education

Feedback in dance refers to the information, judgement, or correction given to a student on how well they perform [30]. Gibbons [30] mentions that dance performance improves faster when students are provided feedback. It is therefore suggested that, adequate and regular feedback may improve the overall execution of dance performance versus when no feedback is given.

2.2.3 Feedback in In-Person Dance Learning

Information from a teacher to a single student, or an entire class might be included in feedback [31]. It can also be a request for a response, allowing information to flow from students to the teacher. It can also create a transactional relationship for the improvement of analytical and critical thinking capabilities [32]. When feedback is purposefully combined with transparent teaching, students can receive thoughtful and reflective feedback from the teacher and learn how to analyze a peer's performance to give and receive different types of feedback. The students can also apply these learned skills to provide feedback to themselves. [31]. In in-person classes, feedback is offered in different ways:

1. *Feedback from Teachers*

It is critical to receive feedback from the teachers after they have observed the learner's performance. Intrinsic feedback is kinesthetic sensory information that occurs naturally within each person as a course of movement. Intrinsic feedback can be received by touch or felt through one's own sense of proprioception (the sensation of where the body is in space, also known as kinesthesia) [31]. It can be heard or seen by students, as in tap dance noises. In short, intrinsic feedback is what is felt by the dancer as they execute a skill or performance. Extrinsic feedback supplied from an outside source can add on to it. Much of the extrinsic feedback provided in a dance class is supplied by the teacher [33].

2. *Mirrors on the Walls* - Mirrors on the walls provide instant visual feedback while dancing. A study was conducted involving dance students to examine how learning and performing dance combinations in mirrored and non-mirrored situations affected student's ability to remember and replicate movement sequences [34]. It was found that there was a marked improvement in scores for dance performed in front of the mirror.

3. *Peer to Peer Feedback* — One of the sources of feedback can be the peers. Sometimes, the provision of teachers' feedback is limited by the class size in dance courses [35]. To solve this problem, peer feedback can be used as an instructional strategy [36]. Researchers have indicated that peer feedback enhances learning by enabling learners to identify their strengths and weaknesses, and by receiving concrete suggestions about how to improve their performance [37].

2.2.4 Feedback in Online Dance Learning

Feedback in online dance learning can be provided by showing recordings of performances by an teacher or by performing in an online meeting setting (using Zoom [11] or Google Hangout [**?**]). Video recording with the help of sophisticated instruments can also assist in generating feedback for performance improvement. Observation and learning from video are important aspects of online dance learning. The students can observe and learn from the teacher video. The process of observing dance performance has also been extensively researched [38] [39]. Maslovat et al. [40] found that due to a mirrored neuron system, observing a skill can induce the body to experience identical neural processes in its motor system as if it were practicing the skill. Similarly, due to the striking similarity in brain activity between observation and execution, self-monitoring has been thought to be superior while seeing another person [40]. Digital video, as an instructional resource, is a form of feedback and self-evaluation. Schwartz et al. [41] is also of the opinion that digital video is a powerful way of learning. According to Schwartz et al [41], saying, seeing, doing, and interacting are the four outputs of the defined learning environment for video use. 'Saying' can provide audible knowledge to the digital video. When the video exhibits things that are presented in a dynamic way, the seeing makes sense. When the video depicts human behaviours, the 'doing' outcome occurs. Finally, when video is used to teach relevant knowledge, it increases student interactivity. Dance students must be involved in the formation of quality conceptions that are similar to those of their teachers (extrinsic feedback from the teacher) and their own conceptions and valued assessments in order to evaluate their own talents caught in film (self-evaluation).

2.2.5 Types of Feedback

In order to design an effective feedback system, it's necessary to know what kind of conditions can impact the dance learning.

The visual, auditory and kinesthetic are the three main systems used to provide feedback to dance learners [33].

1. *Visual Feedback* : Visual feedback may include facial expressions such as a frown or an affirmative smile, hand movements to show the position of the hips in belly dance or a turn in jazz, and a full demonstration [30]. Visual feedback may also involve a written analysis or description, drawing or photograph.

Another example of visual feedback is to show a comparison video between the student and teacher indicating the errors done. In Ballet Hero [42], a wearable technology garment was proposed to help assist dancers' learning by incorporating visual feedback inspired by animation techniques that more directly convey the essential movements of ballet.

2. *Verbal Feedback* : Verbal/auditory feedback may comprise of words such as "Super!", "Drop your elbow just a little" and "Gently, smoothly, with tenderness." Verbal behavior includes vocalizations as well as words, such as singing the rhythmic cue of a waltz, "DUM dee dee, DUM dee dee, dum DAAAH de dum". Other examples include tapping a tap dance combination or clapping a rhythm [30]. Although user testing was not conducted on the system, Essid et al. [43] succeeded in creating an audio-driven virtual dance-teaching assistant that allowed them to create 'augmented tutorial videos highlighting the rhythmic information using, for example, a synthetic dance teacher voice.'

3. *Kinesthetic Feedback* : Apart from verbal and visual, there is another kind of feedback named kinesthetic or haptic feedback. This kind of feedback enables the student to feel corrections, and may include moving the spine into the proper shape for a back arch or having the student in the leader's role in ballroom dance experience the follower's role. Touch may be passive or active. Passive touch is when a body adjustment is physically initiated by someone else, and active when the change is generated by the mover [44]

2.2.6 Feedback Delivery Time in Online Dance Learning

Feedback can be delivered during the performance as well after the performance. Real-time feedback is instantaneous feedback during motion. Post-performance feedback is feedback delivered after the performance. Real-time or immediate feedback can be delivered by means of integrated motion capture and virtual reality (VR) technology [45]. Chan et al [9] integrated motion capture and VR technology to develop a dance training system. In his study [9], student's movements are collected in real time and reproduced by a virtual representation, which is displayed next to the virtual teacher and is made up of cylinders representing body parts. The hue of a cylinder indicates if the body segment is in the proper position. In systems like Super-mirror [46] which is a Kinect interface for ballet dancers, correct poses were recorded previously and a connection was established between prerecorded and live motion-capture data where the color of the animation's legs would change color whenever there is a similarity in movement. Tang at al. [47] used a Progressive-Block Matching algorithm to create a real-time dance game. In addition, Chan at al. [47] developed a method that was similar to the one described above, but focused on doing a full motion analysis of the player's body components in relation to the taught-motion template. Deng at al. [48] created a real-time motion detection system that uses a flexible matching method based on a human body partition indexing technique to determine the end of a move and detect undesired motion. Yang et al. [49] expanded on this work by developing tools for automatically producing dance lessons that adjust to the ability level of the student. Hachimura et al. [50] developed a dance training system that combined motion capture and virtual reality technologies. The movements of a professional dancer are overlayed on the body of the virtual avatar controlled by the user on a head-mounted display. The trainee can see where their body does not intersect with the professional dancer's. In their system, the trainee must do a move while simultaneously seeing their avatar. This might have an impact on performance, and it takes a lot of practise to spot faults. The creators of the 'Folk-dance evaluation' [51], on the other hand, chose to provide the students with post-performance feedback [6]. Post-performance feedback can be delivered by showing interactive graphs, videos of performance or performance statistics after the performance.

2.3 Overview of Motion Capture

Technology has permeated nearly every segment of the modern world, including dance. As dance programme director Marsha Barsky says [6], technology may assist dance instructors in improving their programs, citing the use of 'Annotative tools' (that allow to annotate or describe dance movement or structure during or after the rehearsal or a dance class and have a strong potential for assisting the learning process) [52], and video platforms as examples. In the recent era, teaching and learning of dance has transformed from more traditional approaches to accepting generative ones, improvisation, and creativity [6]. As dance education is embracing more generative and creative approaches, a variety of technological tools including mobile, desktop applications and sophisticated digital environments are being adopted [6]. Motion capture technology can help capture, reproduce, and self-evaluate one's movement and dance performance in the future. Dancers from the Merce Cunningham Dance Company who were filmed using motion capture sensors put directly on their skin acknowledged their delight as per findings from Stanford University by Rosanne Spector [53]. One dancer described the importance of employing technology in the learning process: 'I know what I think my body is doing. But is it really doing that? I don't really know, but I'd like to'.

Motion capture and biomechanical visualizations are examples of technology that can help with comprehension of movement systems and give objective data about one's movement for self-evaluation and further reflection [6]. By analyzing movement from both scientific and aesthetic perspectives, deeper understanding is gained of why people move their muscles and bones in a particular fashion. In other words, when something artistic is being quantified, even if it seems obvious, some knowledge is gained in the process [6]. As a result, recording and analyzing motion may be extremely beneficial in improving dance learning and comprehension.

Nakamura et al. [54] created dance training equipment that includes a timing-vibro device and a robotic screen. The timing-vibro gadget tells users when to move, and the robotic screen moves to display the translation in dance performance for the user. On the screen, a virtual representation of the learner is produced in real time. Users may see and compare their own motions to the expert motion. Hachimura et al. [50] integrated motion capture and virtual reality to develop a dance training system. A head-mounted display shows the professional dancer's motion overlapped with the body of the virtual character controlled by the user. Magnenat et al. [55] presented their advanced technology on digitizing and rendering 3D folk dance where they used the optical motion capture system to capture dance motion from real dancers. Chan et al. made use of [9] an optical motion capture system with cameras and sensors to capture and compare dance movements of teacher and student. This system [9] provides real-time feedback accurately in a small period of time, which is mandatory for real-time feedback to the users. The ballet dance learning environment by Kyan et al. [56] provided instructions through 3D visualizations. The suggested system's architecture is made up of Kinect motion capture and CAVE, gesture detection, and a gesture database. When a novice performs such moves, the system's purpose is to detect a series of predetermined motions performed inside a dance sequence and to identify the occurrence and length of these movements in the dance sequence. The recognition module takes the beginner's performance and extracts occurrences (phases) from it, then compares it to the teacher's (ground-truth) movement. Finally, the technology uses virtual reality to view both the teacher's and student's dance sequences (or single moves). Motion matching has been used for teaching other physical activities as well. Kwon et al. [57] suggested a motion training system for teaching Taekwondo [58] that captures dissimarities between trainer and trainee's movements. This system [57] works in conjunction with visual and bodily sensors. Body sensors are attached to the user's forearm in its design, which may measure two angular values, pitch, and roll of the movement.

Portability is an important aspect to be considered while selecting equipment. Systems that need a lot of heavy or complicated equipment are less portable than others. A Kinect-based system, for example, can be regarded as portable [dance learning games and teaching], but when the system contains complicated augmented reality devices, optical motion capture for the learner, it becomes less so. Super Mirror [46], Webdance [59], and the system [5] for learning Hawaiian Hula are portable systems as they only require screens, kinect sensor and trackers, as is the system presented by Alexiadis et al. [60] in their work for a dancer's performance evaluation. For its purpose, YouMove [61] requires an augmented reality mirror, which is bulky to move conveniently. Same is the situation for Kyan [56] which offers a CAVE virtual reality environment , and the multi dance corpus involves numerous equipment like synchronized audio rigs, numerous cameras, wearable inertial measurement devices and depth sensors to make mobility impossible.

The cost of the instrument is another important factor in equipment selection. In most of the cases cited in different research papers [56] [61], the equipment is expensive. Someone who has just started to learn dance may not be very comfortable to invest too much in buying costly machinery.

In the above mentioned cases, most of the time, the equipment used isn't portable, expensive, or difficult to use. A device containing a camera and an infrared (IR) sensor, or any other system based on magnetic, mechanical, or inertial means, can track the user's motion. It might be a Sony Playstation [62] or a Microsoft Kinect [63]. Initially, in this study, Microsoft Kinect was thought to be used for motion capture. These sensors are well-known in the gaming world. The game is played in real time mode with visualization on a television using one camera and one infrared sensor that tracks the performer. Due to cost, portability and space issues, the plan of using Microsoft Kinect [63] could not be implemented.

Virtual reality applications [9] allow users to immerse themselves in a virtual world and experience a realistic or imaginative virtual world. Users can participate in a variety of scenarios, such as virtual boxing matches [64] or dance with imaginary partners [65]. An optical motion capture system with cameras was the next plan for capturing motion. A suit with markers for motion capture in virtual reality environment [9] was required for the same. But later, it seemed quite impractical for a new learner learning from home.

Head mounted displays can import the dancer to a virtual world where they can dance seeing an expert. Hence, head mounted display was thought of a good option for this project, but it became quite a hindrance while performing full body movements. VR has several other disadvantages as well. Makransky [66] found that used VR contributed to more of mental burden than learning. Wearing head mounted displays for long were also causing physical discomfort like dizziness or 'cyber sickness', which can be experienced as nausea and headaches due to sensory mismatch [67] [68].Hence, the plan of using head mounted display was also discarded.

2.4 Overview on Motion Matching in Online Dance Learning

Dance evaluation is a complicated phenomenon including a range of factors (accuracy, timing, shape, speed, etc.) and hence, it is quite challenging to combine all those different aspects all at once and generate feedback [6]. Due to this reason, most dance games [2] [8] are created to focus on specific components that provide a score-value. In such systems [2] [8], whenever a pupil is given an initial demonstration of a movement and asked to mimic it, the most important feature to be assessed is motion accuracy.

Furthermore, the timing is also an important factor in the estimation of the final score because the student's movement must be in sync with that of the expert [6]. You Move, the Multi-modal dance corpus [61], Chan's reality dance training system, and the works of Alexiadis [60] and Kyan [56] are all systems that attempt to evaluate performance in terms of motion correctness and musical timing (rhythmicality). In 'Super Mirror' [46], accuracy of motion is compared to a movement from an existing database as there is no initial physical demonstration [55]. In this case, the movement is analyzed using a motion recognition algorithm, after which the system extracts the matching move and compares its motion accuracy to the one performed by the student.

Dance motion matching also opens up various new perspectives on giving feedback. Now, feedback can be provided to a student based on different viewpoints like 'Upper-body fluidity'(how graceful the upper body movements are), 'Lower-body fluidity' (how graceful the lower body movements are), 'body balance'(how well the body balance is maintained throughout the performance) and choreography [43]. According to Aristidou et al. [51] , 'most motion analysis methods are based on ad-hoc quantitative indicators, and so do not frequently yield insights on movement aspects of a performance.'

Various studies, such as those by Camurri et al. [69], Fdili et al. [70] and Piana et al. [71], emphasize the significance of movement characteristics rather than movement shape and form in dance performance and learning. Dynamic symmetry is one of the various movement attributes used to evaluate students' dance movement, according to Camurri [69]. Aristidou [51] provides a ‘framework based on Laban Movement Analysis (LMA) [72] concepts that tries to discover style elements in dance motions.’ LMA is a multidisciplinary system that uses Rudolph Laban’s theories to describe, analyse, and document human motions. It includes contributions from anatomy, kinesiology, and psychology.

Aspects considered in Aristidou’s LMA framework are [51] :

a) Motion Accuracy : Motion correctness or motion accuracy refers to how well a student’s posture matches that of the teacher at any given time.

b) Time Frame : When there are a large number of matches between the two postures throughout the course of a whole performance, the motion matching accuracy is presumed to be high.

c) Rhythmicality : Rhythmicality is the expert’s demonstration of movement timing and rhythm which needs to be followed by the user while learning dance.

d) Specific Movement Features : Special movement features such as balance, movement flow, equilibrium, and dynamism has also been referred to in his work.

In the Laban Movement Analysis [51], movement is described as a pattern of change including four distinct elements: body, effort, space, and shape (referred to collectively as BESS). The weights of these four LMA categories are originally equal (25 percent each). The learning platform also allows the user to modify the sensitivity of the system when comparing the motion of the student to the template motions per LMA category.

a) The BODY component is principally responsible for the development of body and body/space relationships. It is responsible for explaining which body parts are moving, which parts are linked, which portions are impacted by others, and what is the sequence of movement between the body parts, as well as general assertions regarding body organization [51].

b) The EFFORT component [51] specifies the movement’s aim as well as its dynamic quality, texture, feeling tone, and how energy is utilized in each action. It is divided into four subcategories:

Space [51] refers to the level of active attention to one’s surroundings.

Weight [51] is a sensory factor that detects physical mass and its gravitational connection.

Time [51] is the body's internal attitude toward the movement's time, not its length.

Flow [51] refers to the movement's consistency. The Effort can be deduced from Head Orientation, deceleration of motion, movement velocity, movement acceleration and jerk.

c) The SHAPE component examines how the body changes shape when moving. It defines the body's static forms, how the body changes as it moves toward a point in space, and how the torso can change shape to facilitate movement in the rest of the body [51].

d) The SPACE component refers to how people move in relation to their surroundings, as well as paths and lines of spatial tension.

Yoshimura et al. [73] captured 3D dance data and utilized an algorithm to extract a characteristic motion from the data which was then applied to dance scenarios performed by a teacher and student. Index variables (indices) [73] created to quantify four common forms were assessed from the extracted movements, with findings showing that they were useful for judging dance competence and individuality. In a dance system of a similar design, Qian et al. [74] created a gesture detection engine. This gesture detection feature recognizes joint angles in ten body parts, including the head, chest, upper arms, forearms, upper legs, and lower legs, but not the hands and feet. The Mahalanobis distance between two sets of joint angles is used to determine how similar two gestures are. Saltate [75] is a wireless prototype system to support beginners of ballroom dancing. It acquires data from force sensors mounted under the dancers' feet and detects steps. It, then, compares their timing to the timing of beats in the music playing, thus detecting mistakes. Alexiadis et al.[60], created a dance training system which visualizes the temporally aligned dance movement of both teacher and students, along with the associated evaluation scores, in a virtual 3D gaming environment. In this system [60], three scores are calculated, which are subsequently combined to produce an overall score. Alexandros et al. [76] propose a dance evaluation system in which a multi-Kinect acquisition system is being utilized, where synchronized skeletal data from each sensor are fused in order to improve the quality of the final skeletal tracking and a new dance scoring methodology is put forward based on fuzzy interference [76].

Table 2.1: An Overview of Different Works

Works of Different Researchers	Motion Analysis		Feedback Type	Portable	Affordable
	Motion Capture Equipment	Motion Matching Method			
Nakamura et al. [54]	timing vibro device, robotic screen, motion capture cameras	Comparison of computer generated images	Real time	No	No
Hachimura et al. [50]	VR headset and MR platform system	Comparison of computer generated images	Real time	Yes	No
Magnenat et al. [55]	Movement trackers placed on body and motion capture cameras	Framewise pose estimation	Real time	No	No
Kyan et al. [56]	CAVE virtual reality environment	Posture identification and gesture recognition	Real time	No	No
Super Mirror [46]	Kinect Camera	Comparing static poses analyzing of knee and hip joints	Real time	Yes	Yes
Web Dance [59]	Motion trackers and camera	3D Animation and web technologies	Real time and post-performance feedback	Yes	No
YouMove [61]	Large displays and Microsoft Kinect	Posture detection from keyframes	Real time	No	No
Aristidou et al. [51]	Movement trackers and screen	Laban Movement Analysis	Real time	Yes	No

Table 2.1 highlights some of the recent motion matching equipment and methods in place in different research. The motion matching technology utilized is quite complicated in several studies when dynamic videos are compared. People from technical background can easily grasp some complex algorithms while someone from humanities may find it difficult as they are not very familiar with complex mathematics. While trying to utilize Laban Movement Analysis concepts and discussing various papers regarding the same, it was realized that specific expertise level is required to understand the multiple concepts of Laban Movement Analysis [51]. A person having a less technical and more artistic outlook may require some time to grasp the concepts of Laban Movement Analysis. Understanding Laban Movement Analysis [51] will be easier for a Laban Movement Analysis dance expert, but it may be difficult for someone from a different background.

2.5 Dance Style

Dance has been an element of human society from the dawn of time, with recorded evidence of dance 30,000 years ago [77]. Dance has also been deeply connected to social life of the people as well the places they reside on. Body movement varies depending on the dance form. Ballet, for example, is a highly technical and codified technique [56]. Movements are precise and rely on many physical attributes to be successful. Some essential theoretical concepts of classical ballet methods are alignment-maintaining verticality of the torso, placement-minimal displacement of the pelvis from the centre position, turnout-maximum external rotation of the lower limbs, and extension-maximum elongation of the lower limbs [78][56]. On the other hand, Bharatnatyam [79] has orchestrated body movements that are synchronized to Carnatic music. Bharatnatyam [79] dances are performed to describe the stories sung using vocal and instrumental music. 'Bharatnatyam is also an expressive dance form that uses semiotics, storytelling, music, culture, religion,mathematics, philosophy, and history and this dance is used to help children of Asian Indian origin to become culturally grounded' [79].

There are some dance styles which require a partner, for example Ballroom dance [80]. Ballroom dance requires two people to participate at any given time, a leader and a follower. Communication between a leader and a follower is essential in ballroom dancing. The leader's job is to begin movement by signaling a follower, and the follower's job is to detect and respond to the leader's stimulus. This necessitates a certain level of cognitive flexibility as well as social intelligence.

Folk-dance is another kind of dance form. This genre of dance is passed down through the generations through social events in addition to being explicitly taught in a studio setting [51]. Folk dancing is particularly 'malleable', as it is altered and adapted over time and across geographical boundaries. Despite the fact that each folk dance has a core set of steps and postures that dominate, folk dancers often enhance and embellish the dance with their own distinctive flair.

Bollywood Dance [81] is the name given to the Indian film dance style. It is difficult to categorize Bollywood in a straightforward manner. Bollywood dance is a unique kind of dance that integrates a variety of forms. This type of dance [81] combines ancient Indian dance styles such as Kathak, Bharatnatyam, Bhangra, and Odissi with western popular dance forms such as bellydance, hip-hop, jazz etc. In the next chapter, the preferred style of dancing for 'Danceshala' is discussed elaborately.

The literature study gives an idea about the different feedback mechanisms, motion capture techniques and dance forms in different systems. To summarize, the following gaps remain in the discussed literature:

1. *Portability* : The recording instrument is not always portable. For example, YouMove [61],CAVE virtual reality environment [56] are not portable.

2. *Affordability* : The recording instrument is not always affordable. For example, sophisticated cameras or bodysuits utilized in some studies ([9],[55])may require additional investments which might not be very feasible sometimes.

3. *Complexity* : The motion capture technology involving bodysuits or sophisticated cameras may appear complicated for a beginner to record movements as they may not be comfortable with

so much technicality. The motion analysis methodology used in the above literature (for example, Laban Movement Analysis [51] and pose estimation [56])is also quite complex and require expertise to understand.

4. *Feedback Delivery issues* : Some of the interfaces like that of Dance games [2] does not give detailed feedback.

5. *Interface issues* : Simple and interactive visual feedback interface for dance learning is missing.

All these gaps mentioned informed the objective of this study which was to ensure the following:

1. *Portability* : The dance recording instrument is easily portable.

2. *Affordability* : The dance recording instrument is affordable.

3. *Complexity* : The motion capture method is simple and motion analysis method is less complicated.

4. *Detailed Feedback* : The interface delivers easily understandable visual feedback in a variety of ways.

5. *Simple and Interactive Interface* : The visual feedback interface is simple ,interactive and can help a beginner to learn dance.

However, the main focus has been given on the interface. The novelty of the research does not lie in finding an unique method in capturing motion and feedback delivery because these methods have already been utilized in different studies. The originality of this research lies in creating a simple and interactive interface with affordable technology which can help beginners in learning dance.

2.6 Summary

All the above mentioned literature study and limitations in the current research methods has supported to reach certain conclusions for different factors helpful for ‘DanceShala’. The design choices and the reason of selecting different factors will be discussed in detail in the next chapter.

Chapter 3

Design Choices

This chapter presents a detailed discussion on the design choices and the reason behind them.

3.1 Visual Feedback

According to Munzner [82], the visual system provides a very high-bandwidth channel to our brains. A significant amount of visual information processing occurs in parallel at the preconscious level. Vanichvasin et al.[83] are also of the opinion that visual communication can transmit information faster and better through its visual elements because it helps process visual information with less language barriers. Auditory feedback is poorly suited for providing overviews of large information spaces compared with visual feedback. In kinesthetic feedback, though touch and kinesthetic perceptual channels are utilized by haptic input and feedback devices, they only cover a small portion of the dynamic range of what humans can perceive [82]. The state of the research shows that there is evidence of the effective use of visual feedback in a number of research settings. Visual feedback has been quite effective in other different scenarios like in sport settings [84] [85], in school settings [86] [87], as well as in mixed settings [88], including some with people with disabilities [89]. In 'DanceShala', the errors in different body parts of the student as compared to the teacher and the severity of the errors at any given time are being pointed out. Hence, here, the visual mode of feedback can provide detailed information about the error. As visual mode of feedback seems to be quite an effective form of delivering information about the inconsistencies, visual mode of communicating feedback is chosen for this study (Table 3.1).

3.2 Feedback Delivery Time

Post-performance feedback is utilized for this study (Table 3.1). The main reason behind this is to not distract the participant by giving instantaneous feedback as the participant is a beginner in dance and hence, is in the nascent stage of learning. Giving real-time feedback too frequently can have negative effect on learning [90]. Continuous real-time feedback appears to enhance performance during practice when the feedback is operating, but it does not contribute to learning and may even degrade learning [91] [92] [93]. Hence, in this study, the visual analysis is provided afterwards. The student sees the reports after their performance, understands where they went wrong , re-practices and then performs again.

3.3 Motion Capture Equipment

In this study, portability and low-cost are the main factors for selecting the equipment for motion capture. Hence, a simple motion capture device which most people possess nowadays and which is used everyday, the mobile phone, is selected. Mobile phones are an indispensable part of our lives now. Apart from being a device for voice and text communications, modern mobile phones or smartphones are extensively utilized for photography and video recording as well. A smartphone requires the usual hardware of a computer system, such as the processor, memory, input devices, output devices, and network interface. They also include a camera, barometer, accelerometer, and digital compass, among other essential sensors [94]. The advanced version of the phone camera found in smartphones allows users to take photos and create videos within a blink of an eye.

The smartphone camera, unlike other types of sensors, has not been extensively used in research. The availability of camera-equipped, processing-capable, inertial sensor-equipped, and reasonably priced mobile phones (a.k.a. smartphones) has piqued the interest of the research community to use them for various study purposes in recent years. The main advantages of mobile phones lies in the affordability and portability. There are disadvantages as well. Capturing the dance motion with a mobile phone will not be very precise as with sophisticated cameras and bodysuits. But a beginner, who just started to learn dance, may not understand the technicalities of using sophisticated cameras or body-suits for motion recording. Therefore, a simple and less expensive device like mobile phone is considered for this research.

Hence, mobile phone camera is used to record dance movements of the learner.3.1) It is as simple as playing a song, and dancing in front of the mobile camera. After recording, an affordable software named RADICAL[13] is used to extract the Biovisional Hierarchy data [95] which is then processed through Blender [96].

3.4 Motion Matching

People interested in learning dance, can hail from any background ranging from science to arts to commerce and they may not be acquainted with the technical aspects involved in motion matching. Therefore, instead of using four different factors as mentioned in Laban Movement Analysis [51], a different and simpler procedure using Euclidean distance [97] is utilized which is discussed in detail in the next chapter.

Joint position, velocity, and angle are three primary characteristics used to measure bio-mechanical differences between two movements. These are three distinct measurements. To find out which one method is the most efficient, an experiment was performed by Chan et al [9]. In that experiment [9],two groups of motion pairs are manually formed: one including motion pairings in which the two movements are identical (Group 1), and the other containing motion pairs in which the two motions are different (Group 2). The motions for the pairings in Group 1 are the exact movements performed by the same or different people. The motions for the pairs in Group 2 are varied movements done by the same or different people. The differences between two movements in each pair are measured and determined using various methods. To calculate the difference between two motion pairs with differing lengths, dynamic time warping is used first to create a mapping between them. The Euclidean distance [97] between joint angles, joint locations, or joint velocities is used to calculate the difference between each matched frame pair. The average distance of all the matched frame pairs is then calculated using the difference between two movements. The result of the experiment showed that among the three metrics, the measure of joint position had the best movement discrimination accuracy [9].

Following the same approach, for this study, difference between joint positions for movement tracking and detecting dissimilarities in motion is considered. The difference between joint positions is studied with the help of a BVH (Biovision Hierarchial data) file (**??**). Deriving the BVH file from recorded student and teacher videos is quite a complex procedure. Initially, the research started with some already available and processed motion capture files available in an online dance database [98] which has a number a publicly accessible digital archives of dances that use state-of-the-art motion capture technologies to record and archive high quality motion data of expert dancers performing traditional dances. In the first step, the BVH files and videos from this database [98] are utilized to test 'DanceShala'. After getting satisfactory visual feedback for these processed files through 'DanceShala', in the next step, BVH files extracted from mobile-recorded videos are utilized for visual feedback generation in 'DanceShala'. The whole motion matching procedure is discussed in detail in chapter 4.

3.5 Dance Movements

To select the appropriate dance style for 'DanceShala', different dance forms are researched. Ballet [56] requires proper technique and body poses, and ballroom dancing [80] requires a partner. As this study is for beginners in dance, the idea is to start with simple dance moves from a dance style so that people from any background can participate and enjoy learning. Bollywood [81] is also quite familiar dance style to me and thus, simple dance movements from the genre of Bollywood 3.1) is selected for this study.

3.6 Selection of Body Parts

In this study, the dissimilarities in movement of the shoulder, arms and legs 3.1) of the student and teacher are considered as shoulder, arms and legs play an essential part to convey expression and meaning in Bollywood dance. There is definitely the scope of adding more body parts but as this is a small scale project created from scratch, shoulders, arms and legs are only considered for now.

3.7 Intended Users for this Study

This application is for any individual who wants to learn Bollywood. The users can be from any background. The only thing required is the interest in learning the art-form. This interface is specially designed for beginners.

Table 3.1: Design choices

Feedback Type	Feedback Time	Motion Capture	Motion Matching	Dance Style	Users	Body Parts Considered
Visual [6]	Post-performance [7]	Mobile Phone	Euclidean Distance between joint Positions [8]	Bollywood	Beginners	Shoulders, arms and legs

3.8 Summary

This section accurately describes the design decisions made while formulating this research. With the assistance of these design choices, 'DanceShala' is created which can exhibit simple and easy-to-understand visual feedback that may enhance the quality of whole dance learning experience of the Bollywood dance form for beginners.

Chapter 4

Methodology

This section will concentrate on the two most important aspects of this study:

1. Dance movement capture (involving Phase 1 and Phase 2 of Figure 4.1)

2. Dance movement analysis (Phase 3 of Figure 4.1).

Here, the methodology as well as the tools and software utilized during the implementation process will be discussed.

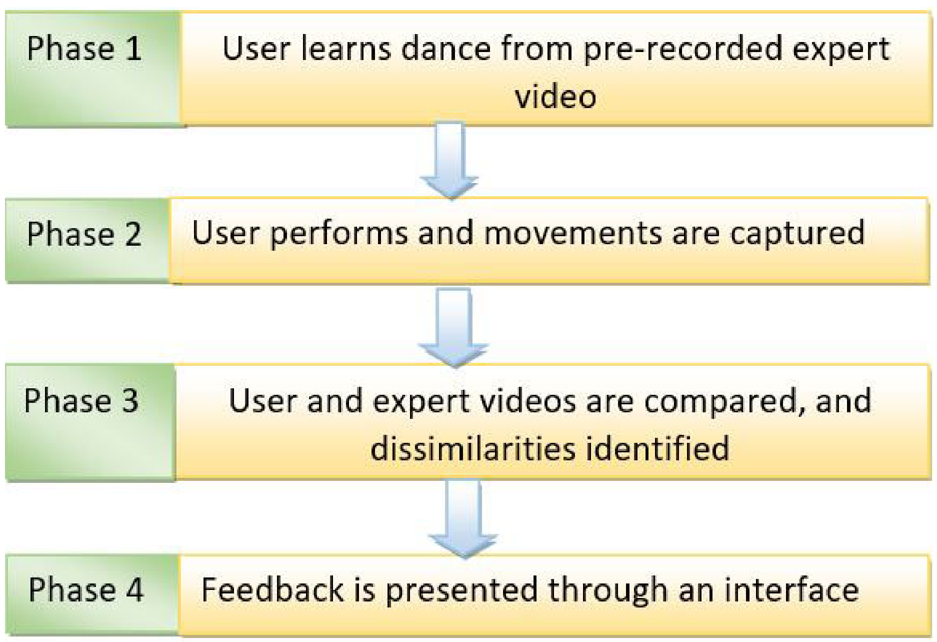


Figure 4.1: The mechanism of the study

4.1 Dance Movement Capture

In this study, simple mobile phones are used for motion capture keeping in mind the portability and cost factor. A motion capture system (the mobile camera in this case) captures a teacher's 'ideal' or ‘perfect’ movement, which is then used to animate an avatar. This avatar is referred to as the teacher’s avatar; it shows the student how to move and is also used to occasionally compare with the student’s movement. In the following sub-sections, some important points regarding the video recording are discussed.

4.1.1 Teacher Video Recording

An instructional teacher video is recorded first. The purpose of this video is two-fold. Firstly, it is used by a student to learn dance steps. Secondly, it is used as a reference when comparing with the student’s recorded video to provide feedback. This video shows some dance movements on a particular audio track and is recorded with the help of a mobile camera. The video can be recorded anywhere; time and place do not matter. However, few important things need to be taken care of while recording the video.

1. *Record on any Phone*: A very expensive high-end mobile phone is not required. The video can be recorded on a mid-range mobile phone having a good camera quality. We obtained satisfactory results on an android phone recording in 720p at the rate of 30 frames per second, although higher resolutions and frame rates can also be used but will add to the processing time.

2. *Visible Movements*: The teacher should be facing the camera; the movements should be clear and pronounced so that the student can easily discern them when viewing the video.

3. *Simple Movements*: The choice of dance steps is up to the teacher, but in this initial stage, the focus is on simple movements that span a short duration and are easy to follow.

4. *Good Audio Quality*: The teacher is free to choose any audio track that goes well with the dance. However, the lyrics and beats need to be clearly audible as they are important in Bollywood dance. It can be very difficult to learn dance from an instructional video without these elements.

5. *Background*: The video needs to be processed for further analysis and it is best to have a simple uncluttered background.

4.1.2 Observation of Teacher Video for Learning

Students watch and learn from the teacher's video. They observe and mimic the teacher's movements to learn the choreography. Using the video and audio as guides, students practice the movements at their own pace in a self-guided manner.

4.1.3 Student Video Recording

After observing and learning from the teacher video, the student performs in front of a camera and records a video of their performance. The student can perform repeatedly and check their performance accuracy each time. This can be an iterative process until the student fully learns the movements and performs with minimum error. For the student video, the audio does not need to be recorded.

4.2 Dance Movement Analysis

The captured videos consist of raw video frame data. In order to extract movement information from these videos for comparison purposes, some processing steps are needed. These processing steps are explained in detail below.

4.2.1 Methodology for Movement Matching from Video

For movement tracking and recognizing dissimilarities in motion, the differences between joint locations are analyzed. For this, the BVH (Biovision hierarchial data) file format is utilized, and simple Euclidean [97] distances between joint positions of the student and the teacher are calculated to quantify the dissimilarity.

4.2.2 BVH File Format

The BVH file format is commonly used to store motion capture data. A BVH file is divided into two sections: a header section that explains the skeleton's hierarchy and beginning position, and a data portion that provides the motion data. The keyword 'HIERARCHY' appears at the start of the header section. The next line begins with the keyword 'ROOT', followed by the name of the hierarchy's root segment to be specified. It is permissible to establish another hierarchy after this one is specified ; this one, too, is designated by the term 'ROOT'. While a BVH file can include several skeletons, it is common to just designate one per file, and this is the approach we follow as well. The specific hierarchy used in our work is shown in Figure 4.2. More details about the BVH file format are elaborated in Appendix **??**.

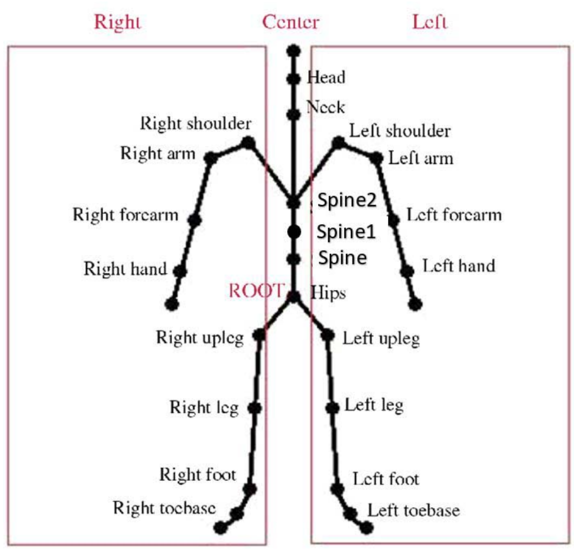


Figure 4.2: BVH Skeletal Hierarchy

4.2.3 Obtaining BVH Data from Video

Extracting 3D motion data from 2D video frames is a computationally intensive inverse problem. State of the art solutions rely on artificial intelligence (AI) and machine learning techniques. RADICAL AI [13] is a commercial web-based service that employs an advanced computer vision system for recognising and recreating 3D human motion from 2D video data. To do so, several fields of research are combined in a unique way, including computer vision, deep learning, and anthropometry. In this project, the recorded videos are fed into this software which then produces as output, frame-by-frame BVH data for both teacher and student videos.

4.2.4 Analyzing the BVH File

Matrix arithmetic is utilised to compare motion data points between the teacher and the student. The BVH file stores position and orientation information for each joint. This information can be transformed to appropriate rotation and translation matrices which when combined, represent a transformation of the local reference frame associated with the joint relative to the root. Note that in our case, there is no scaling information stored in the BVH file. By multiplying the rotation matrices for each of the different axes in the order they occur in the hierarchy part of the file, the rotation matrix, R, may be simply constructed. Here, an example of a bone's channel description can be considered (Figure A.1):

```
CHANNELS3 Xrotation Yrotation Zrotation
```

The matrices R_x, R_y and R_z that rotate about the three axes are:

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{pmatrix},\ R_y = \begin{pmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{pmatrix},\ \text{and}\ R_z = \begin{pmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.1}$$

where $\theta_x = \texttt{Xrotation}$, $\theta_y = \texttt{Yrotation}$ and $\theta_z = \texttt{Zrotation}$ are the rotation Euler angles corresponding to the three axes. The composite right-to-left rotation matrix, R, based on the independent rotation matrices about each axis is given by

$$R = R_x \cdot R_y \cdot R_z. \tag{4.2}$$

The final rotation matrix R is also a 3×3 matrix.

An individual bone's motion is made up of translation, rotation, and scale components (depending on the channels provided for the bone), which may be combined to generate an overall transform in homogeneous coordinates. The sequence in which these distinct transformations are combined to provide the entire 4×4 transformation matrix M will always take the form shown in the following equation where S, R, and T are the scale, rotation, and translation matrices, respectively.

$$M = T \cdot R \cdot S. \tag{4.3}$$

Since we don't have any scale transformation, the combined transformation matrix takes a simpler form. The translation components are merely the first three cells of the fourth column and the rotational components take up the top left 3×3 cells. The composite transformation matrix is constructed using a homogeneous coordinate system, and is given by

$$M = \begin{pmatrix} R_{11} & R_{12} & R_{13} & V_x \\ R_{21} & R_{22} & R_{23} & V_y \\ R_{31} & R_{32} & R_{33} & V_z \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.4}$$

The root bone is usually the only one with per-frame translation data, but each bone has a base offset that must be added to the local matrix stack. As a result, the total of a bone's base position and frame translation data are represented by V_x, V_y, and V_z which are the translation components of the bone's base position relative to the parent with respect to the three coordinate axes.

The data is given in a hierarchical fashion in most motion capture file formats, thus the matrix in Equation 4.4 only yields the local change of a bone. A bone's local transformation explains its orientation inside its local coordinate system, which is susceptible to the local orientations of its parents. The global coordinates for each bone origin may be determined using the following global transformation matrix which takes into consideration the location transformation matrices of each bone.

$$M_n^{\mathsf{global}} = \prod_{i=0}^{n} M_i^{\mathsf{local}}. \tag{4.5}$$

Here, n is the current bone whose parent bone is $n-1$, and $n = 0$ is the bone at the root of the hierarchy. The position of the bone can be determined from its origin in its local coordinate system by using the offset information in the hierarchy part of the file. For example, the end-points b_0 and b_1 of the right foot bone (see Figure 4.2) are given by:

$$b_0 = M_{\mathsf{hips}} M_{\mathsf{RightUpLeg}} M_{\mathsf{RightLeg}} M_{\mathsf{RightFoot}} \cdot [0,0,0,1]^T, \text{ and} \tag{4.6}$$

$$b_1 = M_{\mathsf{hips}} M_{\mathsf{RightUpLeg}} M_{\mathsf{RightLeg}} M_{\mathsf{RightFoot}} \cdot b \tag{4.7}$$

respectively, where b contains the local coordinates of the right foot bone (in the coordinate system of the right foot), M_{hips} is the transformation matrix of the hips, $M_{\mathsf{RightUpLeg}}$ is the transformation matrix of the right upper leg, M_{RightLeg} is the transformation matrix of the right leg, and $M_{\mathsf{RightFoot}}$ is the transformation matrix of the right foot. The vector $[0,0,0,1]^{\mathrm{T}}$ in Equation 4.6 denotes the right foot's local origin, which is translated into its global coordinates by Equation 4.6.

4.2.5 Calculating the Euclidean Distance between Bones

In order to calculate the Euclidean distances [97] between corresponding bones of the teacher and the student, the following procedure is utilized.

1. At the beginning of the recording, both the teacher and the student stand in a reference T-pose in front of the camera so that the full body can be seen including the head, neck, hands, legs, fingers, and toes. In computer animation, a T-pose, also known as a bind pose or reference pose, is a default pose for a 3D model's skeleton before it is animated (Figure 4.3).

Figure 4.3: Standing in a T-pose.

Figure 4.4: Animated T-pose

2. Both the teacher and student videos are processed by Radical AI [13] which yields BVH files that use the reference T-pose at the beginning of the videos to infer the skeletal hierarchy. The BVH files generated by Radical AI for teacher and student follow the same hierarchy. An animated avatar is then created with the assistance of MakeHuman software [99] which is a free and open source 3D computer graphics middleware designed for the prototyping of photo-realistic humanoids. The animated avatar is superimposed on the BVH skeleton with the help of Blender [96], which is an open-source 3D computer graphics software tool-set used for creating animated films, visual effects, art, 3D-printed models, motion graphics and interactive 3D applications. The animated avatar is deployed to give a realistic touch to the teacher and student videos.

3. From the BVH files, per-frame joint positions of both the teacher and the student are calculated by traversing through the entire hierarchy as explained in Sec. 4.2.4. In order to determine the Euclidean distance [97] between two corresponding joint position, the Frobenius norm is used.

The Frobenius norm, sometimes also called the Euclidean norm, is a matrix norm defined as the square root of the sum of the absolute squares of its elements. It can be written as

$$\|A\| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}^2}, \tag{4.8}$$

where $\|\cdot\|$ denotes the Frobenius norm, and A is an $m \times n$ matrix.

Suppose that the transformation matrices associated with two corresponding bone positions are T_1 (teacher) and T_2 (student). If we apply these transformations to the vector y, we obtain the vectors y_1 and y_2 according to:

$$y_1 = T_1 \cdot y \text{ and } y_2 = T_2 \cdot y. \tag{4.9}$$

If the two transformation matrices T_1 and T_2 are relatively close to each other, then the corresponding vectors y_1 and y_2 should also be close. Let T_d be the matrix that transforms T_1 to T_2, i.e.

$$T_d \cdot T_1 = T_2. \tag{4.10}$$

The matrix T_d also transforms y_1 to y_2. In particular,

$$T_d \cdot y_1 = T_2 \cdot T_1^{-1} y_1 = T_2 \cdot y = y_2. \tag{4.11}$$

From the two transformation matrices T_1 and T_2, we can determine T_d according to:

$$T_d = T_2 \cdot T_1^{-1}. \tag{4.12}$$

The Frobenius norm of T_d gives us the desired Euclidean distance between the joint positions y_1 and y_2.

4. Finally, the error is normalized to a value between between 0 and 100. Any error beyond a value of 1500 is mapped to 100 and errors below 1500 are linearly mapped to a value between 0 and 100.

The aforementioned procedure is used to calculate the error in the positions of the shoulders, hands and legs. The errors or the dissimilarities between the movements of the teacher and the student are then conveyed to the student by means of an interface which will be discussed in detail in the next chapter.

4.3 Summary

The setup primarily focuses on how the motion is captured and how the motion is analyzed to find the differences between the teacher motion and the student/learner motion. The movements are recorded on a mobile camera and then are processed through RADICAL [13] to generate BVH files. The BVH files are then analyzed to find out the difference in movements. These dissimilarities in motion help to generate feedback which is later visualized through the interface.

Chapter 5

Visual Interface

The visual interface is an important component of our research. In this section, the design aspects and methodology of the visual feedback interface in 'DanceShala' are discussed in addition to some related works which have influenced the interface design.

Before creating a successful interface, human variables must be addressed. These considerations might include things like memory and age. For example, a sophisticated user interface is not appropriate for children according to Cen et al. [100]. Elderly users [101] require additional text help in order to identify software components and avoid probable mistakes. Hence, before creating the interface, the designer should decide on what kind of persons or age-group is being targeted. 'DanceShala' is created for assisting beginners to learn dance. Hence, the interface is kept relatively simple and straightforward.

5.1 Structuring the Design

An effective usable design consists of three different design aspects [102]: 1) Interface Design, 2) Information Design, and 3) Interaction Design. These three design qualities should be incorporated and combined in order to make a flawless design (Figure 5.1).

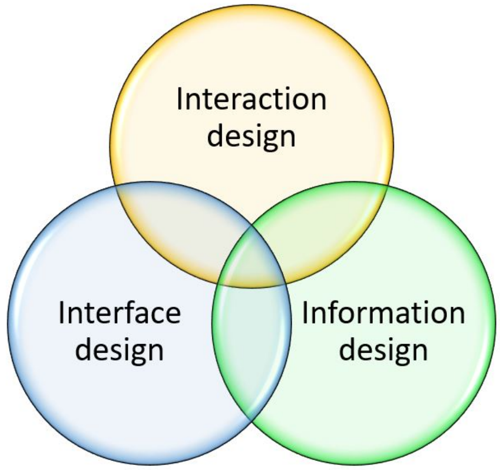


Figure 5.1: Usable design

1. *Interface Design*

An interface is the medium of communication with the user and thereby, delivering important and valid information in the process. Colors, objects, buttons, menus, and icons should all be utilised. It should be simple to navigate. A poor user interface will result in a worthless system [103][104].

2. *Information Design*

The information should be presented in such a way that the user can simply determine who is the intended recipient, including text, graphical items, photographs, and diagrams. There must be information on the link between values and changes in information [102]. The systems must be able to anticipate what the user is doing and send messages that are relevant to the situation.

3. *Interaction Design*

When creating an interface system, factors such as the user's age, culture, and background should be considered, as well as the user's computer skills and capacity to adapt to the system. When looking for information, the user should be able to navigate and get help quickly. There should be no tension or strain on the user [105] [106] [107].

This interface is for the beginners who are trying to learn dance. Hence, considering the beginner user base, all these three factors, interface design, interaction design and information design have been incorporated in design of the visual interface in 'DanceShala'.

1. *Interface Design in 'DanceShala' Visual Interface* : The main objective of 'DanceShala' is to create a suitable dance learning platform suitable for beginners. Hence, the interface of 'DanceShala' is designed in an easily understandable and consistent way by effectively utilizing simple objects like videos and graphs along with tooltips, buttons, menus and icons. These elements are organized in a web-based dashboard as shown in Figure 5.2.

2. *Information Design in 'DanceShala' Visual Interface* : In 'DanceShala', visual feedback information about student's dance movements is displayed through different visual elements like score, upper and lower body analysis, graphs as well as videos of dance movements of both teacher and student. Precise information of error like error time and magnitude of error of individual body parts is also conveyed to make it easier for the beginners to understand.

3. *Interaction Design in 'DanceShala' Visual Interface* : As 'DanceShala' encompasses a varied range of users specially beginners in dance learning, hence, human aspect have been duly considered. Attention is given to improve the interactions between the user and the interface so that the beginners can find it useful to navigate between different elements and analyze their movement errors. Interactions between different elements like graphs and videos are added to create an interactive and informative interface. Simple and understandable graphs are utilized to make the learning process uncomplicated for beginners without any extra strain and tension.

5.2 The Design

The interface (Figure 5.2) includes the following components:

1. Score chart

2. Student and teacher videos and video controls

3. Upper body and lower body analysis

4. Main graph

5. Error graph

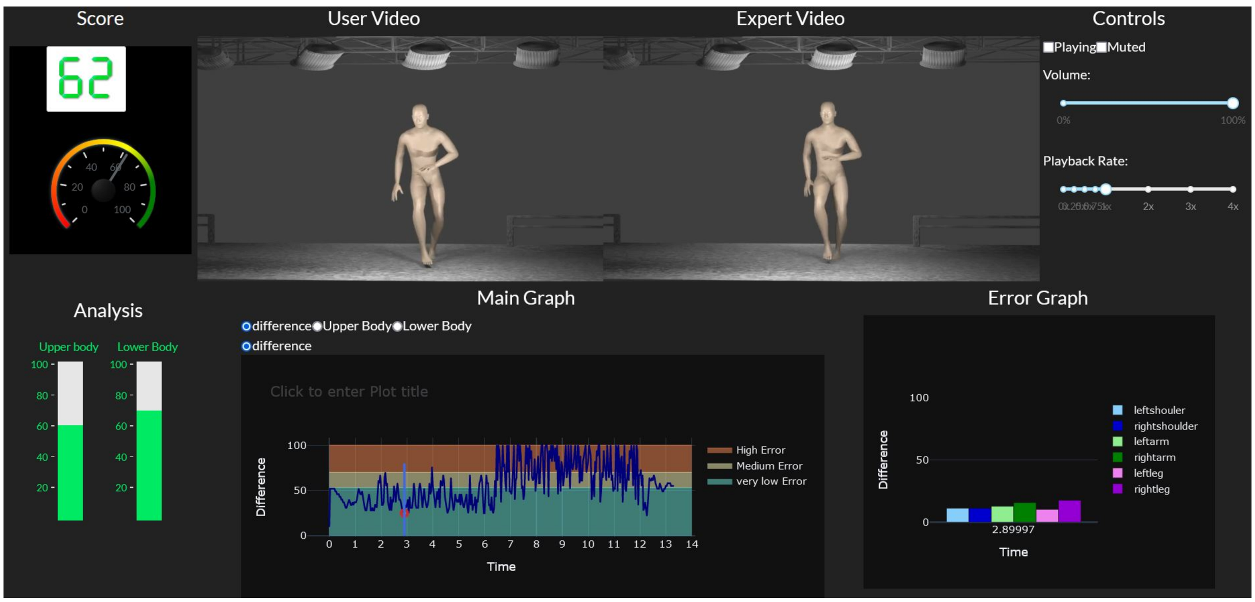


Figure 5.2: Visual Feedback Interface

5.2.1 Score - Overall Evaluation

Score is an initial point of evaluation of a performance in this visual feedback interface. Score is a way of giving some feedback which can serve as one of the important catalysts for student learning. A score has the ability to indicate the level of proficiency of the student who is performing.

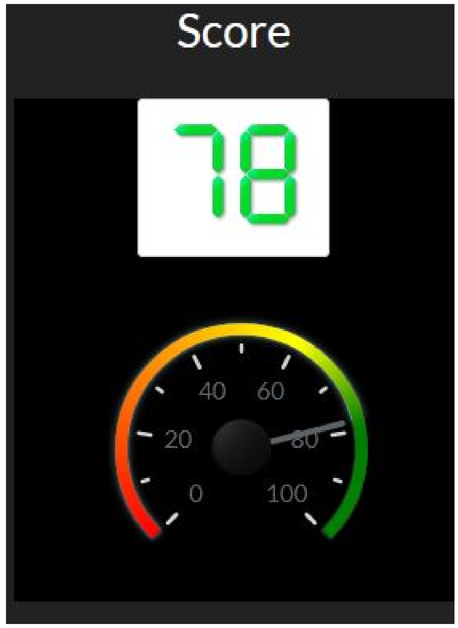


Figure 5.3: Score

Figure 5.3 is the demonstration of the score as found in the top corner of the left hand side. The score section is divided into two parts:

1. *The Numerical Score* : The numerical score, which is shown in the upper part of the score section in Figure 5.3, indicates a number which gives an initial idea of the performance to the dancer – the higher the score, the better the performance. The score ranges from 0 to 100. As any evaluation process, there should be a standard based on which the student's performance is gauged. Here the teacher's video serves as the standard and the perfect ideal video. The total average error is average of all normalised individual body parts error. All the error values lie between 0 and 100 (the procedure of calculating the individual body part errors is described in Section 4.2.5). For this project, as stated previously, left shoulder, left arm, left leg, right shoulder, right arm and right leg are considered and the overall error is given by

$$E = \frac{1}{6}\left(E_{\mathsf{LeftHand}} + E_{\mathsf{RightHand}} + E_{\mathsf{LeftLeg}} + E_{\mathsf{RightLeg}} + E_{\mathsf{LeftShoulder}} + E_{\mathsf{RightShoulder}}\right). \tag{5.1}$$

Then the score S is calculated as:

$$S = 100 - E. \tag{5.2}$$

In layman terms, the score can be thought of as an accuracy indicator (i.e how accurately the performer has performed with respect to the teacher).

2. *The Gauge chart*: A gauge chart is also displayed below the numerical score (Figure 5.3) to provide a better understanding to the learners as to how much they have learnt and how well they have performed. This chart resembles a car speedometer. A diverging red-green colour map is used to indicate the score. The colour gradually transitions from green to red as the score goes from 100 to 0. It's the most green at 100 as green symbolizes success. It turns to yellow midway which indicates average score and needs some improvement. The colour gradually starts changing into red as the score reduces further signifying that the student needs to practise more to improve their performance.

5.2.2 Teacher and Student Videos

Video is a continuous medium with a plethora of information that may be utilised to figure out what's going on in a scene. As a dance video is a dynamic video consisting of several frames, a single frame does not give enough information to find out the differences between teacher and student movements. Several frames must be evaluated to figure out the difference of movements. In 'DanceShala', teacher and student videos are juxtaposed for comparative evaluation.

1. *Performance Videos*: The teacher video is an important component of the interface. The animated avatar corresponding to the teacher's performance is shown on the right (Figure 5.4). The students may compare their own performance — shown on the left — with the teacher video and see the differences in a frame-by-frame manner.

The videos are the main elements that users are attracted to. Therefore, they are kept on top of the interface. Additionally, the teacher and student videos are juxtaposed for easy comparison. Moreover, the same simple stage environment is used for the two avatars, so that the student is not distracted while comparing their performance with the teacher's. Furthermore, spot lighting and a static camera are used to identify the movements easily.

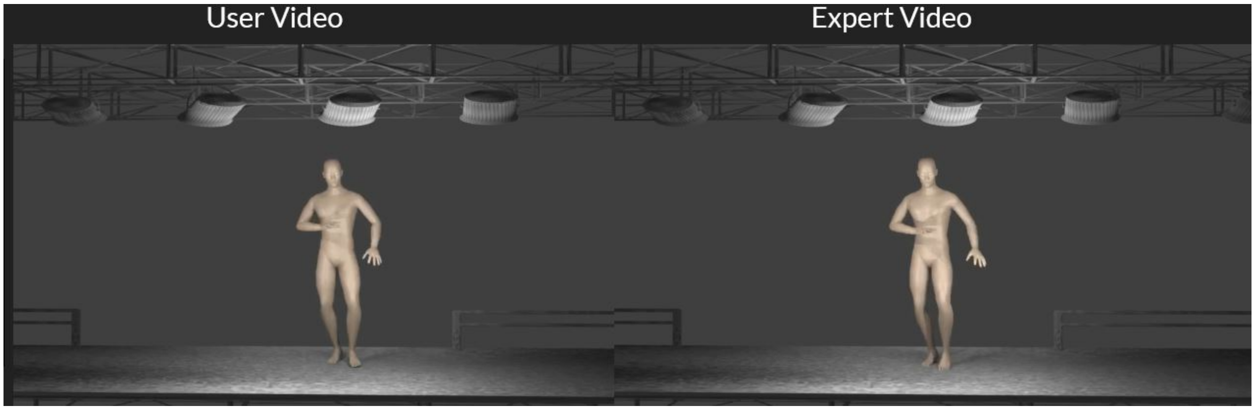


Figure 5.4: Student and teacher videos in the interface

2. *Video Controls*

The same controls are used to navigate both videos at the same time (Figure 5.5); i.e. the videos are affected synchronously via these controls.

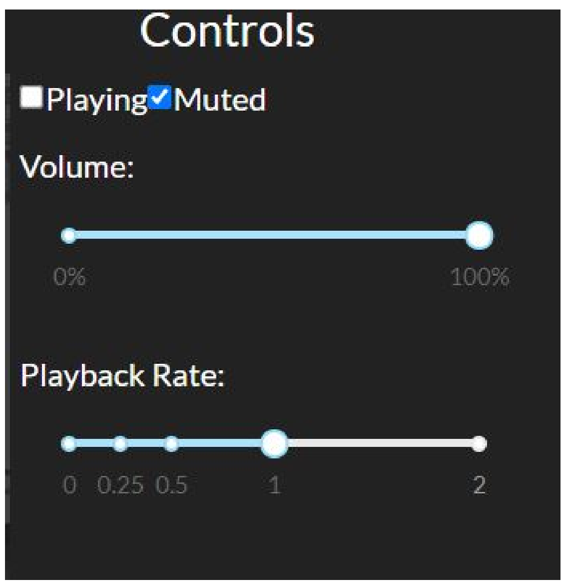


Figure 5.5: Video controls

The video control section is placed on the top right corner of the interface. The video control section has the following options:

(a) *Play and Mute checkboxes*: The videos can be paused or played as per student's convenience by toggling the 'Playing' checkbox (Figure 5.5). Similarly, a 'Muted' checkbox is provided to mute the audio track. Muting is a good option for those students who want to observe the videos to identify their improvement points diligently without the audio distractions.

(b) *Volume slider*: A slider is placed in the next row to control the volume. Music is an important factor in dance videos; different learners may prefer different volume levels. The slider adjusts the music volume to any level from 0 percent to 100 percent (Figure 5.5).

(c) *Playback Rate*: Playback rate, as the name implies; indicates the playback rate of the videos. Different learners learn at different rates; some are fast learners, and some prefer learning at a slow pace. For some learners, the movements can be really tricky at times. This feature is specially designed for those who want to observe and learn the moves slowly to perfection. With the help of the playback rate slider, the student and teacher videos can be played at different speeds as shown in Figure 5.5.

5.2.3 Upper and Lower Body Analysis

This feature of the interface informs a student about how their upper body and lower body performed.

Figure 5.6: analysis

1. *Upper Body*: The upper body analysis shows whether the upper body parts of the student were in sync with that of the teacher with respect to the choreography learnt. The upper body parts, as the name signifies, refer to the body parts above the hips, particularly, the left shoulder, right shoulder, left arm and right arm. Hence, the upper body score in this case refers to the accumulative score of these body parts only. In Figure 5.6, the upper body score is around 70. Other upper body parts can be considered in future enhancements.

2. *Lower Body*: The lower body analysis, in the same way as the upper body, indicates whether the lower body parts of the student were in sync with that of the teacher with respect to the choreography learnt. In this study, the lower body specifically refers to the left and right legs. Therefore, the lower body score refers to the accumulative score of the left and right legs only. In Figure 5.6, the lower body score is around 85. There is scope of adding other lower body parts in future enhancements.

5.2.4 Main Graph

The main graph demonstrates a clear and detailed view of the performance. It shows the difference between the student's movement and the teacher's movement at any time of the performance. The *x*-axis indicates the time in seconds and the *y*-axis indicates the difference of movement between the student and the teacher. The difference is in the range of 0 to 100 as calculated in Section 4.2.5. The 'time' varies as per the total time of the performance. Figure 5.7 shows an example of a 14 second video.

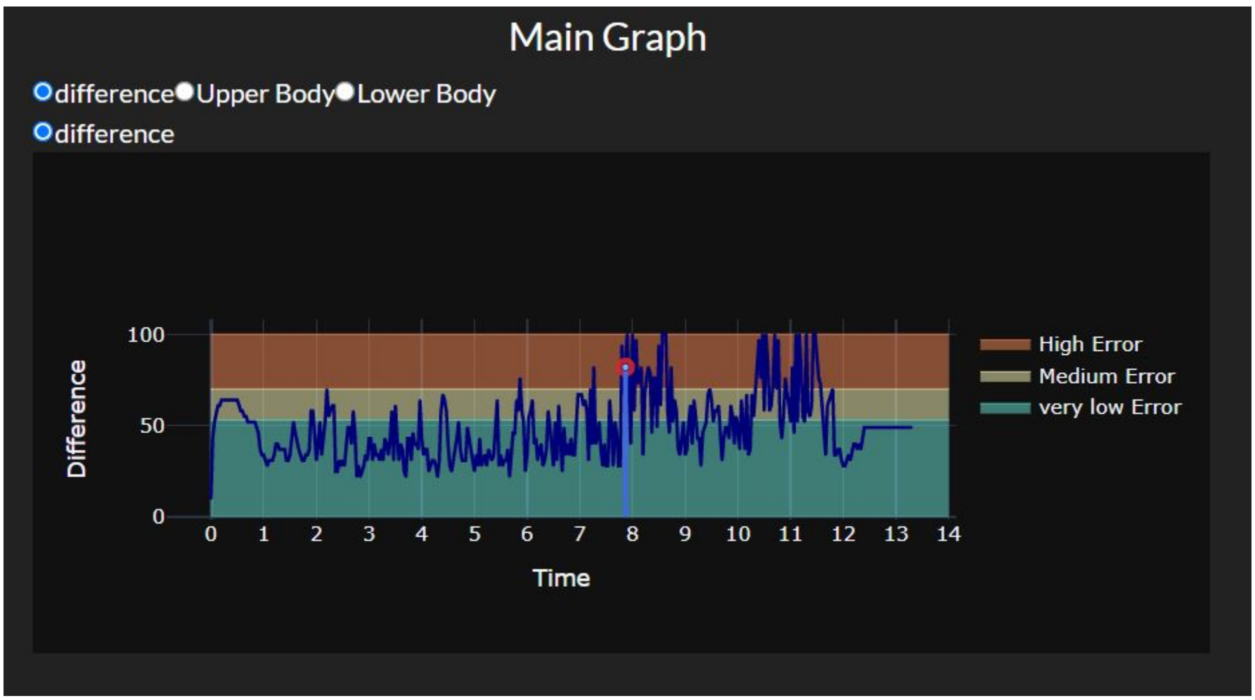


Figure 5.7: Main graph

The main graph view also incorporates the following features:

1. *Difference/Error Zones*: In this study, no sophisticated instruments are used for recording. Mobile recorded videos have different variations as they are not recorded in closed and controlled environments. So multiple recordings and testing with different dance videos were conducted to develop qualitative standards to assess the performances. This video analysis revealed that there

is little discernible difference between the teacher and student performances when the error is between 0 and 50, some discernible difference when the error is between 50 to 70, and large discernible differences above 70. Therefore, 50 and 70 are taken to be the thresholds for medium and high errors and the area behind the error graph is divided into three zones (Figure 5.7). Anything below 50 is in the *low error* zone, 50 to 70 is the *medium error* zone, and anything above 70 is in the *high error* zone. For easy visual identification, the error zones are also colour coded.

2. *Click functionality*

The advantage of an interactive interface is that the student can control the different features to have an intriguing dance learning experience. Interactivity is an important aspect of the interface in 'DanceShala' which is achieved by the click functionality. If any point on the graph is clicked, three things occurs (Figure 5.8).

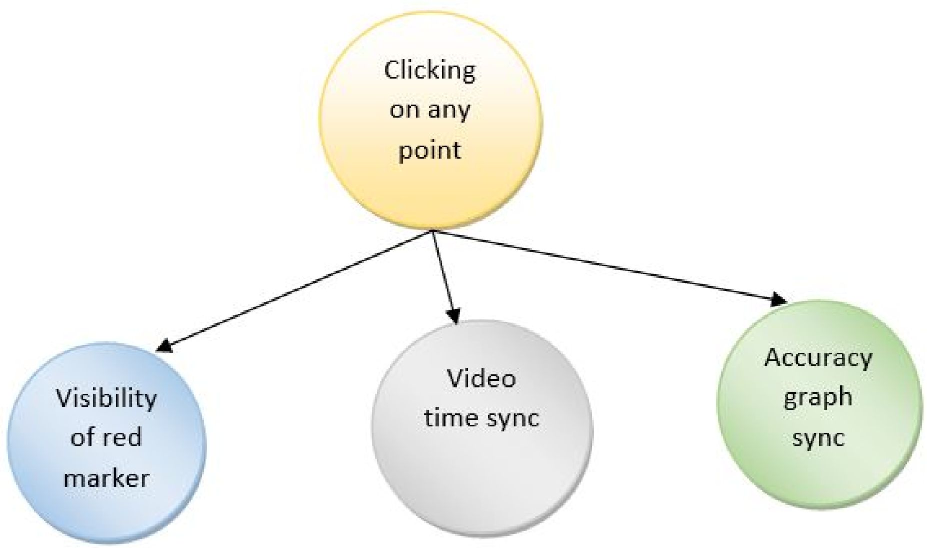


Figure 5.8: Click events

(a) A red marker is created at that point giving better visibility to the student (Figure 5.7).

(b) The click also controls the student and teacher video timings. When the user clicks on any point in the graph, the student and the teacher videos also come to that same point. This is really a helpful feature when combined with video playback speed. If the students want to observe the difference of movements at a specific time and learn from it, they can simply click on that point and once the video comes to that point, they can play the video at 0.25 or 0.5 speed for better understanding of the movement.

(c) The Accuracy graph (Figure 5.11) is also manipulated by the click. When the student clicks on any point, the accuracy graph shows the difference in movements of left shoulder, right shoulder, left arm, right arm, left leg and right leg at that point of time. Visualizing the difference of movements in so many ways really make dance learning easy for the student.

3. *Radio-buttons*: The main graph (Figure 5.7) has three radio-buttons, namely ‘Difference’, ‘Upper Body’ and ‘Lower Body’.

(a) *‘Difference’ radio-button*: When the difference radio-button is clicked, it shows the overall difference between student and teacher movement at each time-frame (Figure 5.7). When any point in the difference graph is clicked, the student and teacher video moves to that point of time and the accuracy graph (Figure 5.11) also shows the difference of movements between the student and the teacher at that time.

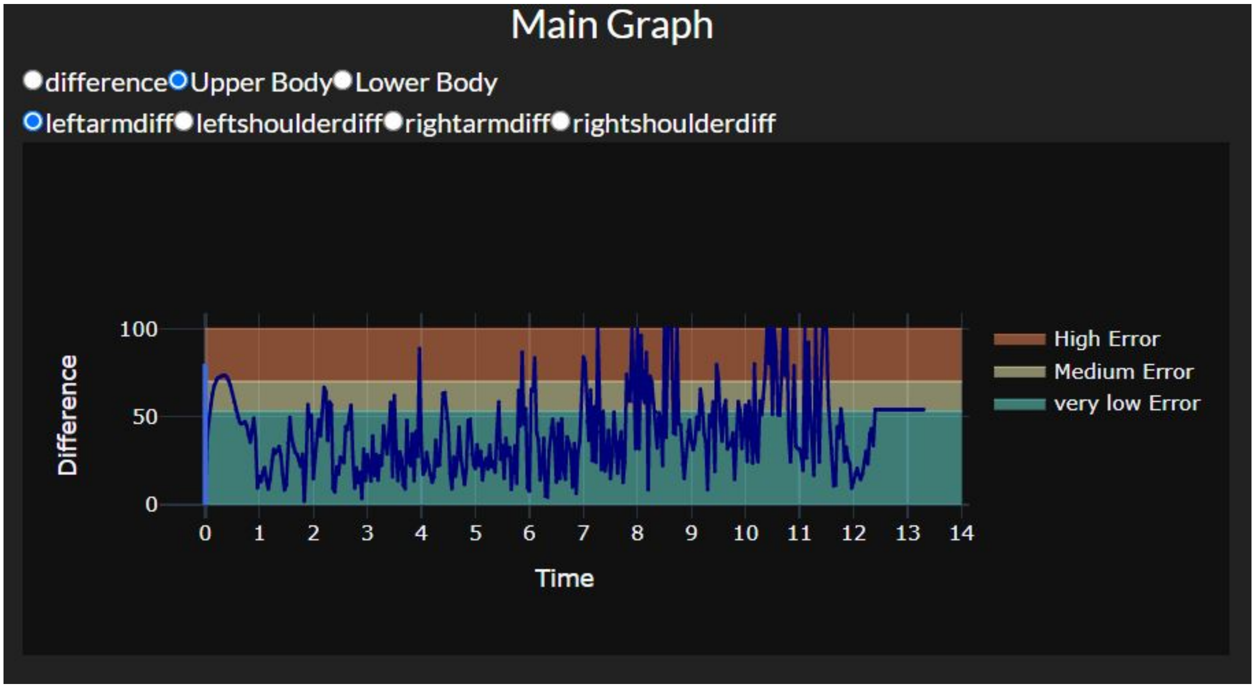


Figure 5.9: left-arm difference graph

(b) *'Upper Body' radio-button :* The 'Upper Body' radio-button once clicked, gives the option of four more radio-buttons (Figure 5.9): 'leftarmdiff', 'leftshoulderdiff', 'rightarmdiff' and 'rightshoulderdiff', which, when clicked, displays graphs indicating the difference between the left arm, left shoulder, right arm and right shoulder movements of the student and the teacher dancer respectively throughout the performance.

The click functionality works in the same manner for these buttons as well. The student can navigate to any specific part of the video by clicking on that specific time on the graph.

(c) *'Lower Body' radio-button :* The 'Lower Body' radio-button once clicked, gives the option of two more radio-buttons: leftlegdiff and rightlegdiff (Figure 5.10). These radio buttons display graphs indicating the difference between the left leg and right leg movements of the student and the teacher dancer throughout the performance. The click functionality works similarly for these graphs as well.

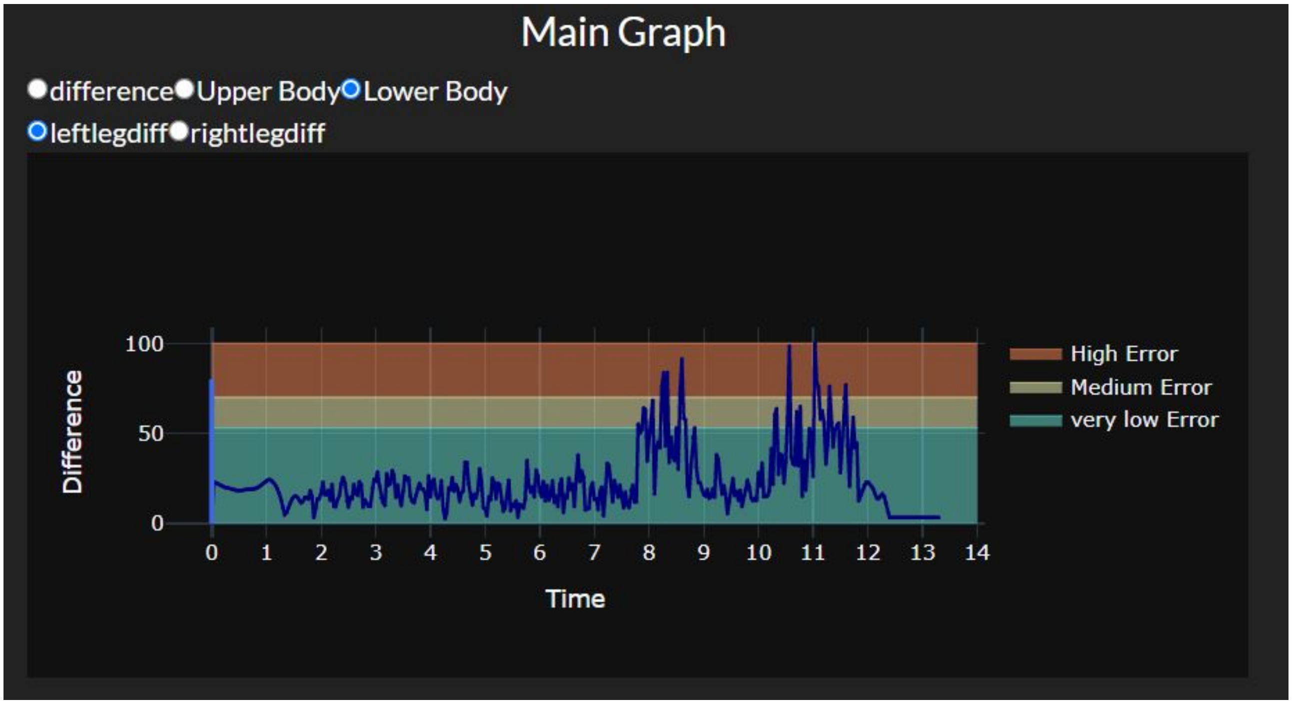


Figure 5.10: left-leg difference graph

5.2.5 The Accuracy Graph

The accuracy graph is the bar graph which manifests the accuracy of the different body parts at any particular time.

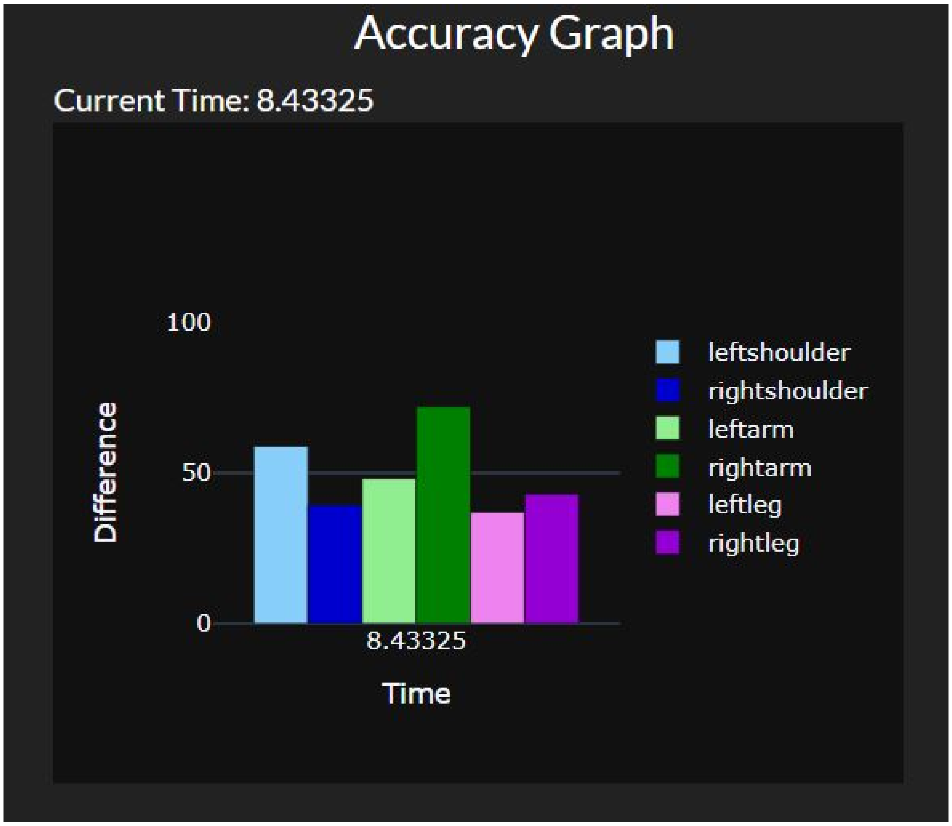


Figure 5.11: Accuracy graph

The *x*-axis for a particular time is a categorical axis while the *y*-axis indicates the movement difference between the student and the teacher. There are six color-coded bars (categories) which indicate how each body part is performing. For example, the deep blue bar indicates the **right shoulder** movement difference between teacher and student. In Figure 5.11, the bar is below 50 at around 8.43 seconds, which means the student and teacher movements of right shoulder at that time are almost in sync.

5.3 Summary

When it comes to 'DanceShala', the design and precision of the interface is crucial for displaying the visual feedback of dance performance. This section highlights the characteristics of the interface in detail. The interface has been designed with various elements which provide a broader view as well as a detailed view. ‘Score’ and ‘Upper and Lower Body Analysis’ give an overall idea of the performance, while the graphs and the videos are meant for in-depth analysis. The interface design is kept simple and straightforward and provides relevant information regarding the error time, error intensity and error position. Interactions between different elements like graphs and videos make the interface more intuitive and informative.

Chapter 6

Survey

This study is a proof-of-concept/prototype of a dance-learning platform which would help beginners to learn dance. In the last step, a survey is conducted with some experienced dancers to receive feedback about the interface for future modifications. In this chapter, the survey details and the feedback received are explained.

6.1 Relevance of Survey

A survey is a way of obtaining information from a group of individuals. There are many methods in HCI for obtaining feedback about research like Ethnographic interviews, log data analysis, card sorts, usability studies and surveys [108]. In this research, this survey is treated as an opinion survey [109] as this survey is designed to get opinion about the effectiveness of the platform and improvement plans after showing and explaining the interface to the experienced dancers. In this research, surveys are used to receive feedback regarding 'DanceShala' owing to the following factors:

1. *Attitude :* Through this survey, the experienced dancers' views about 'DanceShala' as a dance learning platform are assessed.

2. *Successful Fulfillment of Objective :* This survey is conducted to determine an experienced dancer's opinion about whether the objective of teaching dance to beginners can be fulfilled.

3. *Feedback on Improvement :* It is mandatory to know about feedback about 'DanceShala' for future modifications.

Surveys can be done offline as well as online. Online survey was selected for this research study owing to the following factors:

1. simple access to vast geographic areas (including international reach),
2. ease of data entry and analysis,
3. low cost of administration, and
4. timeliness and flexibity.

6.2 Survey Participants

Though ‘Danceshala’ is for beginners to learn dance, we wanted some feedback regarding the dance learning platform. To gather feedback regarding the effectiveness and future improvement plans for ‘Danceshala’, a survey is conducted with experienced dancers as an experienced dancer will likely be able to ascertain the effectiveness of a dance learning interface owing to the experience. The recruitment criteria for the survey participants were as follows:

1. The participant needs to be above 18 years.
2. The participant should have at least one year of experience in teaching or learning dance.

Recruitment poster (Figure B.2) was published in the University of Calgary website and emails (Figure B.1) were sent to University of Calgary attendees which contained information about the study. 25 people responded showing interest in participating. Elaborate description of the survey procedure and consent form were sent to them. Of the 25 people contacted, 15 provided informed consent. The survey was then conducted with the 15 participants who provided informed consent. A small incentive of $10 was provided to every participant who completed the survey.

6.3 Different Scenarios for Visualization

Three different scenarios are created which represents various use-cases of ‘DanceShala’. I have created the teacher and student videos by recording myself. One video recording represents the teacher while three other student recordings mimic different scenarios as described below.

1) In the first scenario (Figure 6.1), the student followed the teacher in an almost perfect manner with very little or no error. As a result, the score was high and the graphs showed the errors in low severity level.

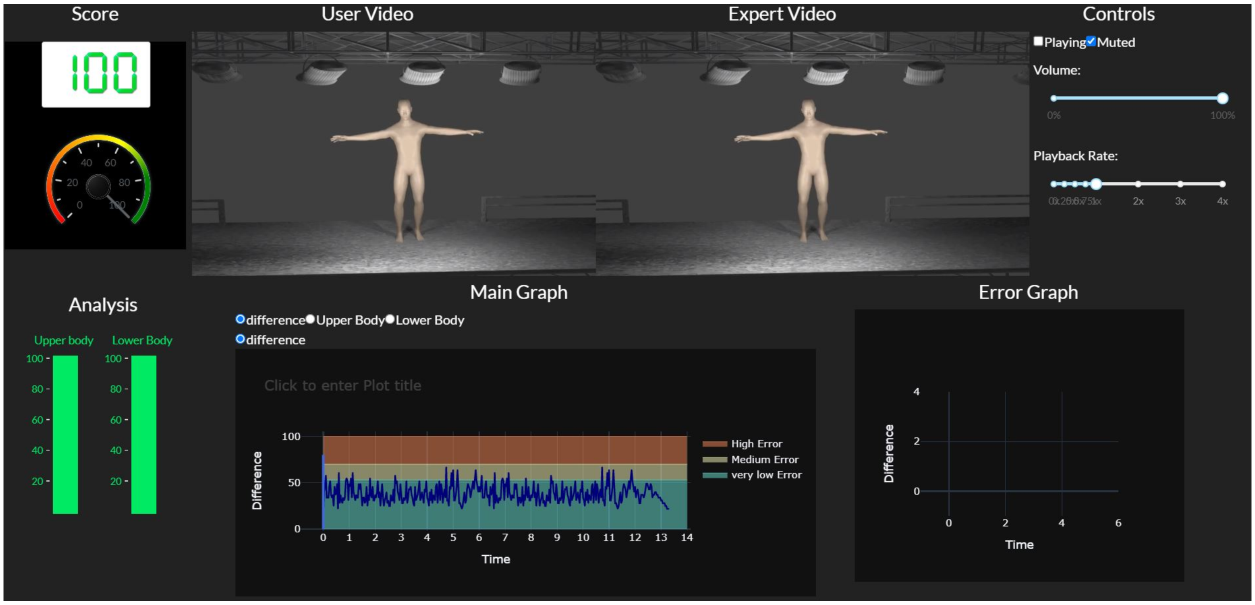


Figure 6.1: First Scenario where student follows the teacher religiously with very little or no error

2) In the second scenario (Figure 6.2), the student made small mistakes as compared to the teacher. As a result, the overall score dropped a bit; the errors were in low severity level most of the time.

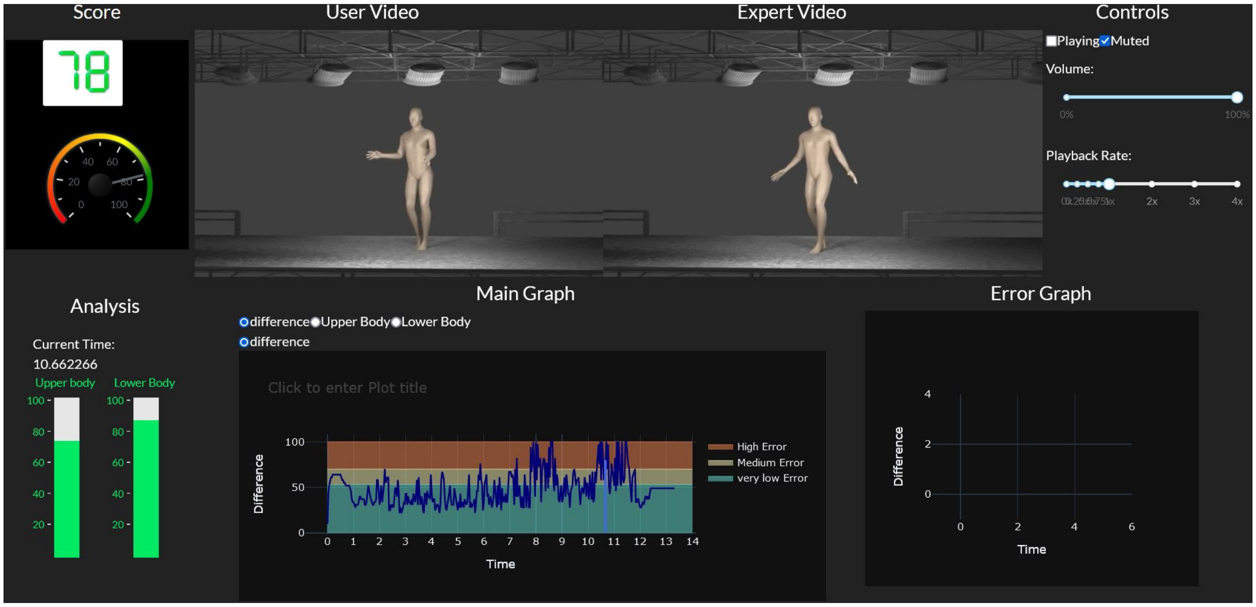


Figure 6.2: Second scenario where student makes small mistakes

3) In the third scenario (Figure 6.3), the student made prominent mistakes. Hence, the score was lower than the above two scenarios and the errors were prominent in the graphs as well.

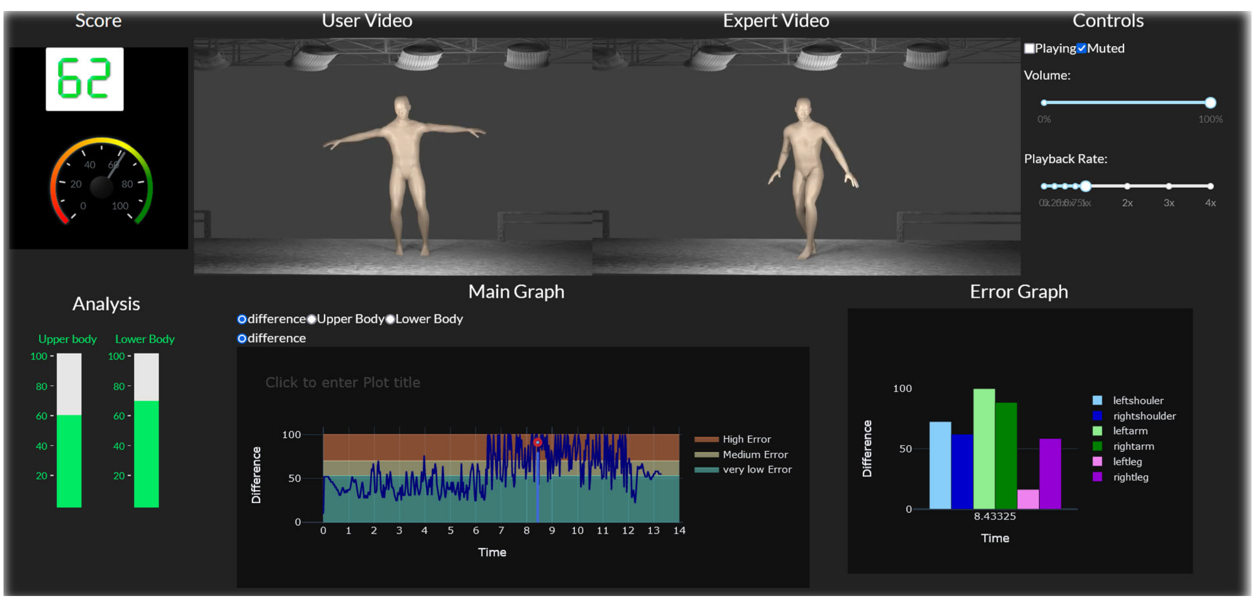


Figure 6.3: Second Scenario where student makes major mistakes

6.4 Survey Procedure

The survey administration procedure followed the steps below (Figure 6.4):

1. First, the web interface was hosted online through Amazon Web Services so that the participants can access it on their own through a link.

2. Ethics approval was sought and received from the Conjoint Faculties Research Ethics Board (CFREB), University of Calgary.

3. The survey was created in Qualtrics so that the participants can easily access it via a link.

4. A recruitment email (Figure B.1) was sent and a poster (Figure B.2) was also published in University of Calgary research participation website for reaching more people.

5. The informed consent process is one of the central components of the ethical conduct of research with human subjects. It is the process of informing potential research participants on the major components of a study and what their involvement would entail. In this research, informed consent was sought from all the participants before participating in the survey.

6. After receiving consent from the participants, they were given an online demonstration of the 'DanceShala' via Zoom on a one-on-one basis. They were shown and explained different features of the application elaborately. It was an interactive session and their questions regarding the same were also answered.

7. Afterwards, online survey link and hosted interface links were sent to the participants through email.

8. The participants used the link to explore the different elements of the interface again. They assessed the interface and filled the online survey.

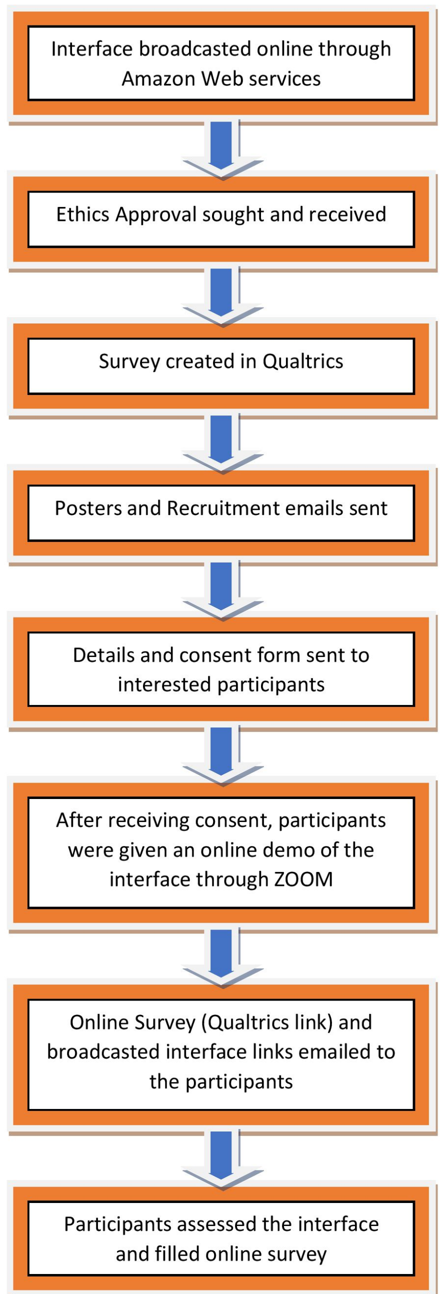


Figure 6.4: Survey procedure

6.5 Survey Questions

Survey questions are an important part of the research. The effectiveness and future development of 'DanceShala' can be gauged with the help of answers to these questions. The detailed survey questions can be found in Appendix B.2.

6.6 Survey Response

The survey responses are as follows:

1. In the first and second question, the participants were asked about their age range and their experience in dance. A total of 15 dancers participated in our survey; 11 were between 18-29 years old, 3 were between 30-40 years old, and 1 participant was in the 40-59 years range (Figure B.5). 11 participants out of 15 had more than 10 years of experience in dance.

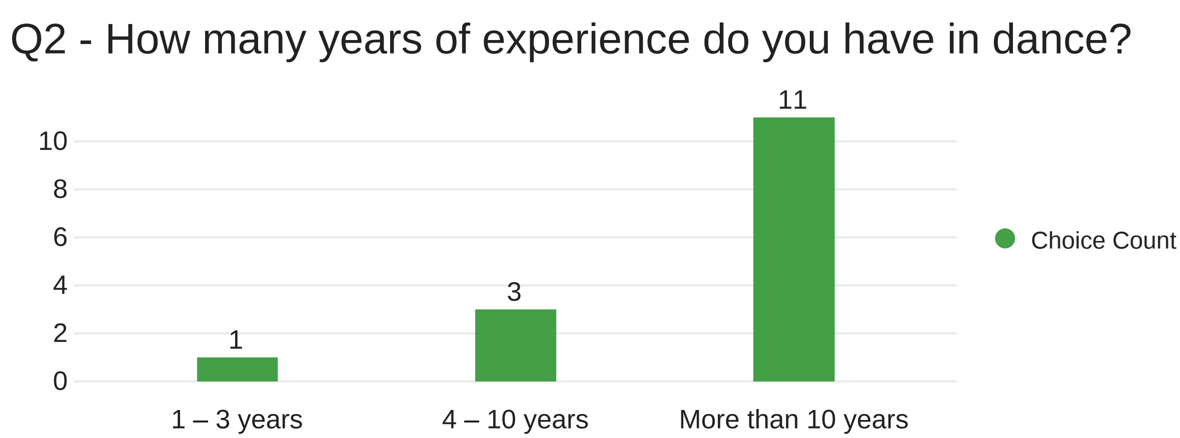


Figure 6.5: Dance experience of the participants in the survey

2. The third question dealt with the understandability of the information visualized in the interface. 8 participants were of the opinion that the information is somewhat easy to understand and rest 7 participants informed that it is easy to understand (Figure 6.6).

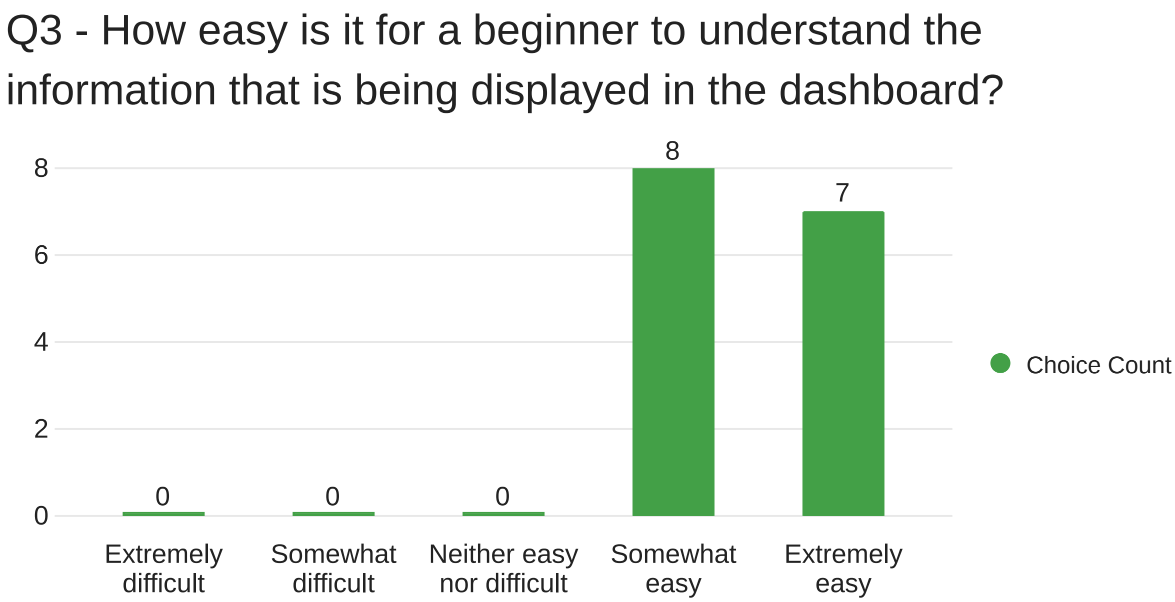


Figure 6.6: How easy is it for a beginner to understand the information that is being displayed in the dashboard?

3. In the fourth question, the participants were asked if they agree with the statement that 'DanceShala' will help a beginner to learn dance. 9 of the participants strongly agreed to the statement while none disagreed 6.7).

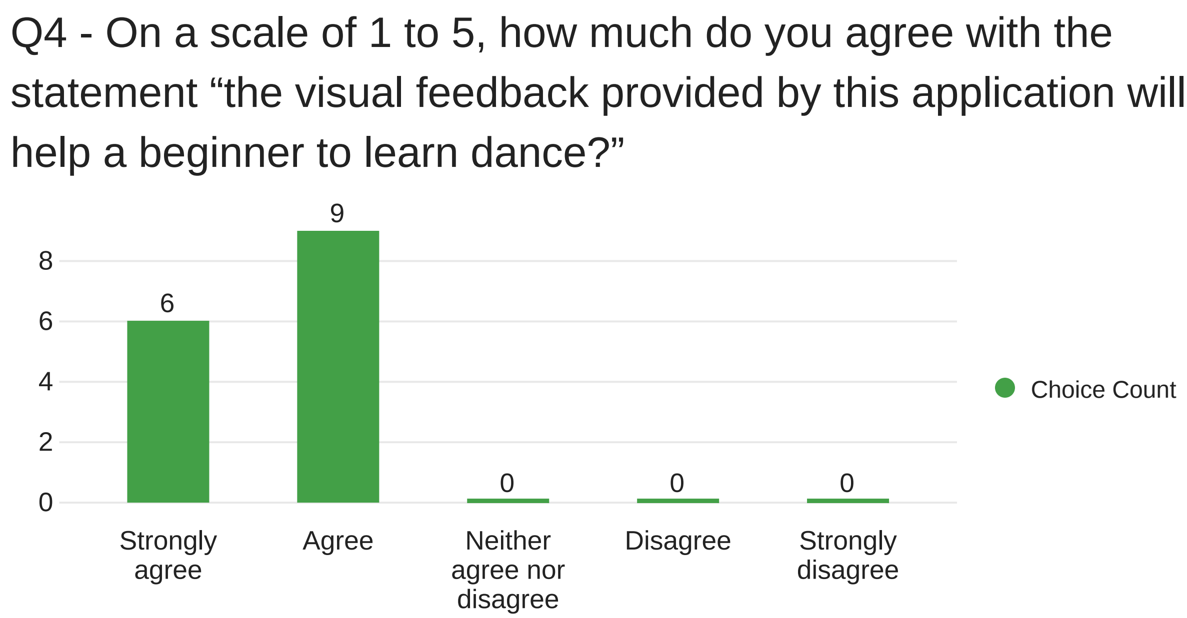


Figure 6.7: How much do you agree with the statement "the visual feedback provided by this application will help a beginner to learn dance?"

4. The fifth question dealt with the usability of the controls provided in 'DanceShala'. Only 1 among 15 participants was of the opinion that its difficult for them to control the different elements of the interface (Figure 6.9).

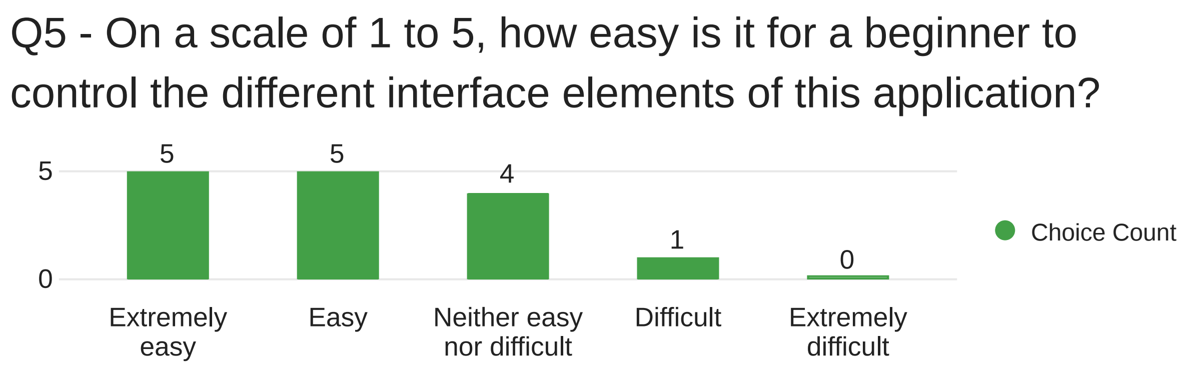


Figure 6.8: How easy is it for a beginner to control the different interface elements of this application?

5. In the sixth question, the dancers were asked if the level of interaction provided in 'Dance-Shala' is useful for a beginner to learn dance. 8 participants found it very useful and 7 participants found it useful for learning dance.

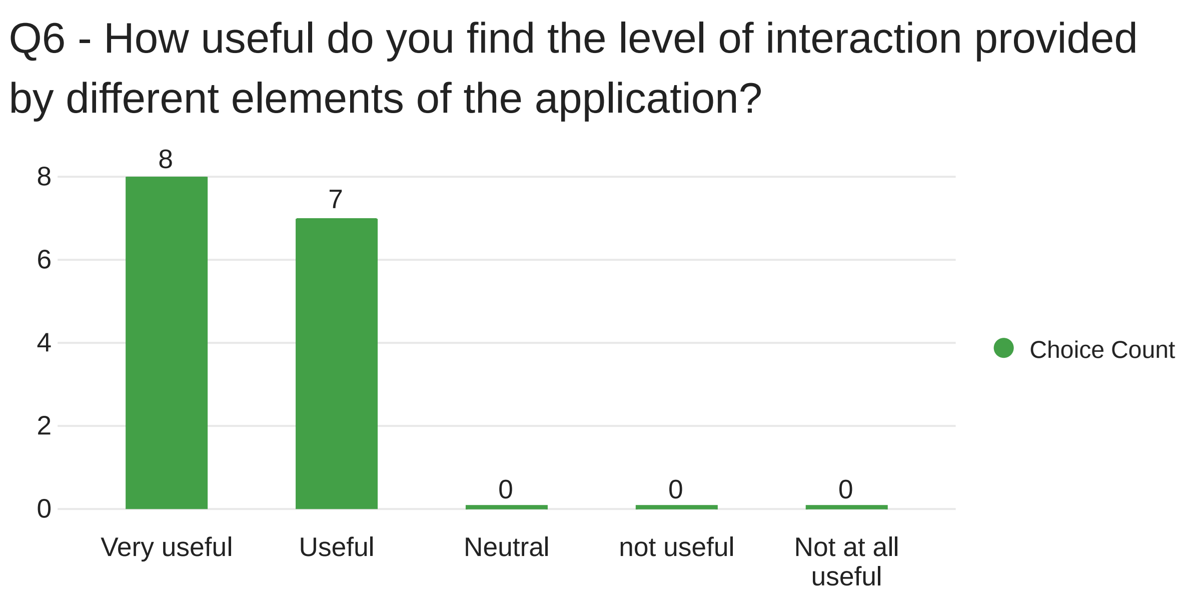


Figure 6.9: How useful is the level of interaction provided by different elements of the application?

6. In the seventh question, the participants were asked to select the most important visual

element for a beginner (Figure 6.10). 4 participants said main graph is important , 4 were of the opinion that error bar chart/accuracy graph is beneficial, 3 of the participants agreed that videos will help a beginner and rest 4 preferred the score and analysis view. Hence, it can be inferred that none of the elements in the interface is non-essential.

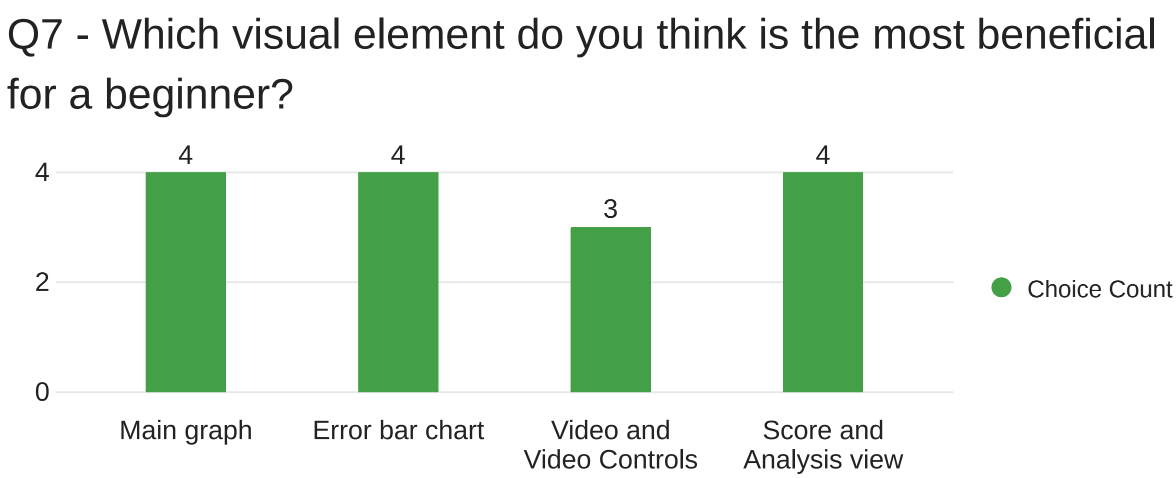


Figure 6.10: Which visual element do you think is most beneficial for a beginner?

7. Finally, in the last question, the participants were requested to give suggestions on improvement. These suggestions are discussed in more detail in the next section.

6.7 Survey Feedback

11 responses on improvements were received. For better understanding, the participants who responded are defined as P1 to P11. P1 stated that the interface is a great tool for learning dance. According to P2, that there is nothing to improve. As some dancers might struggle with the interpretation of the numbers, P3 suggested to write a qualitative description of the errors and discrepancies. P4 commented: 'I believe this experience for a beginner dancer who truly wants to perfect a certain movement can prove to be very useful. I would consider changing the 'error bar' wording to little less demeaning word. Perhaps 'improvement bar'? Overall, this is an interesting project

that has potential and I'm excited to see where it goes.' As it is a minor change, this suggestion is taken into consideration and the name of the error bar chart or error graph is changed to 'accuracy graph'.

P5 and P6 suggested to do some more improvements on the look and feel of the interface. P7 said, 'I honestly can't think of anything that needs to be fixed. I think this is an incredible software for dancers, especially dance teachers to evaluate and analyze their students dancing. I personally found it really satisfying seeing a score describing how well the student danced because that was a clear indication of the quality of your performance'.

P8 is of the opinion that dance should be divided into sections. For example, if the dance is 30 seconds long, the dance should be divided into two 15 second sections to allow the student to practice on a smaller amount of choreography before attempting the entire dance from start to finish. According to them, this will be useful because it allows the student to focus on a specific part and not have to think about completing the entire dance right away.

The video aspect is very important according to another participant of the survey. According to P9, 'a beginner learning to dance most likely would not be concerned with differentiating the errors between the upper and lower body as they are becoming functionally bodily aware; with all their limbs creating mental pathways. The video aspect is ESSENTIAL!'. P10 stated that it would be beneficial to have some highlights when the student does wrong movements. P11 is of the opinion that it is a great interface. All the suggestions will be considered as improvements in the later upgrades of the application.

6.8 Summary and Limitations

The survey gave an idea about a dancer's reactions towards 'Danceshala'. The survey had different restrictions as well. The survey was done in the middle of the pandemic, so in-person demonstration could not be provided to the participant. For the testing purpose as well, external candidates could not be invited as physical distancing was necessary. Hence, I had to create my own videos both as teacher and student creating different scenarios. Much diverse scenarios could have been assessed if more people were recorded.

The survey was conducted with 15 people. There was a set of 8 questions determining the effectiveness of the interface. Hence, it can be considered as a small scale survey. Additional participants and questions could have provided a much improved and diverse response. In addition to that, testing ‘DanceShala’ with a wider group of audience will attract lots of different people to try to learn dance through the application.

Though video demonstrations of ‘DanceShala’ were provided to the participants for assessing the interface properly, still an in-person demonstration could have helped a participant to gain more insight to ‘DanceShala’ resulting in better and constructive feedback. However, the suggestions provided by the participants provide insight into areas where ‘DanceShala’ is lacking and can be improved in future versions.

Chapter 7

Conclusion and Future Work

Dance training takes time and money, and face-to-face instruction is a great method to receive training and feedback. But what happens if a student is unable to find a teacher due to time, monetary or location constraints? The prospects of this research in such scenarios are enticing. The created visual feedback interface can make dance learning available to students all around the world, and they can do it without the concern for physical barriers between teachers and dancers. 'DanceShala' welcomes individuals from all walks of life. On the design side, the interface is kept basic and straightforward, so that it doesn't require considerable inquiry to uncover any specific functionality, thereby decreasing the student's stress. The features are developed with the idea that the system is for learners, and hence, 'DanceShala' is not complicated. Too many technical computer terminology and names are avoided, as some of the students might not be familiar with them.

Feedback has always played a significant role in the dancer's performance and learning. In this study, a bottom-up approach is proposed. Here one of the most important elements of learning dance, feedback, is being studied extensively, and is then interconnected to future studies to form a complete system of dance-learning. The whole study can be summarized in five major steps. In the first step, different feedback methods are studied and the reason of selecting visual feedback is explained. In the second step, a thorough analysis is done on the different equipment used for dance capture in various scenarios and clarifications are given on the usage of a simple video-capture equipment, a simple mobile camera for this study. In the third step, motion-capture and motion-analysis process are explained. In the fourth step, a visual feedback interface is created which contains results of the final motion analysis between the teacher and the student. In the fifth step, a survey was conducted with some dancers and dance-teachers to collect feedback on the effectiveness of 'DanceShala' and everyone almost agreed that 'DanceShala' will really help beginners to learn dance.

There are various limitations. Firstly, the videos are one of the most important components because the video analysis helps to draw a conclusion on the similarity or dissimilarity of the movements. But it should be kept in mind that the video must be recorded with a very specific starting posture. If this starting posture is interrupted (for example, if the hands or legs are not visible properly or are in different start position than prescribed), the results may be inaccurate. A person from a non-technical background may not understand the methodology and may do the recording incorrectly. This will not only create a flawed video, but will also affect the other features of the interface like the main graph, accuracy graph, score and analysis. Secondly, the videos are not recorded with sophisticated instruments. As the videos are recorded using phone cameras, there are certain disruptions and the results are not always perfect. Finally, as pointed out by one of the participants in the study, graphs can be daunting for a person who is not very knowledgeable about technological aspects. Hence, a student needs to have at least some basic knowledge about the technicalities before using the interface to their fullest benefit.

This is an initial study and there are lots of scope for future improvements. For future studies, firstly, more qualitative aspects will be considered in 'DanceShala' as has been pointed out by one of the respondents. There will be more descriptive texts mentioning dissimilarities in each body part which can help a beginner to analyze their dance perfectly. Secondly, the visual feedback display will be improved upon. Mirrors help the dancers to get feedback instantly in a dance studio. In short, mirrors provide instantaneous or real-time feedback. In future research, real-time feedback will be incorporated in addition to the already existing post performance feedback. The real-time feedback would be optional and users can choose if they want to get real-time feedback. Thirdly, some small yet sophisticated instrument will be utilized for capturing the teacher and student movements. The recording is currently done with a mobile camera, but a more sophisticated equipment like high definition cameras or bodysuits with reflective markers will only reduce the minor glitches and help to deliver a more persistent feedback which will result in improvement of the dance learning. Fourthly, in this study, the videos considered have a short 15 second duration. In later research, videos of longer duration will be incorporated. Fifthly, there is an abundance of dance styles that are not considered in the current version. In forthcoming studies, 'DanceShala' can be seen as a universal dance learning platform providing dance learning for a variety of dance forms. The student can select to learn from a wide range of dance styles. Many folk dance styles can also be revived in this way.

This research fulfils the short term goal of exhibiting the capability of providing visual dance feedback thereby supporting dance learning remotely. This study also shows a possible path on how technology may be aligned to support teachers in assisting students. The originality of the method suggests that performance can still be improved with technology-aided visual output. In 'DanceShala', alternative notions of human-machine interaction are explored, particularly for beginner dancers, thereby adapting to present and future generations who have grown up with digital technology. By amalgamating art and technology, 'DanceShala' can be considered as a step-forward in the world of dance education.

Bibliography

[1] Sylvie Fortin. Tomorrow's dance and health partnership: the need for a holistic view. *Research in dance education*, 19(2):152–166, 2018.

[2] Jacob Smith. I can see tomorrow in your dance: A study of dance dance revolution and music video games. *Journal of Popular Music Studies*, 16(1):58–84, 2004.

[3] Sylvie Fortin. Tomorrow's dance and health partnership: the need for a holistic view. *Research in Dance Education*, 19(2):152–166, 2018.

[4] Katerina El Raheb, Akrivi Katifori, and Yannis E Ioannidis. Hci challenges in dance education. *EAI Endorsed Trans. Ambient Syst.*, 3(9):e7, 2016.

[5] Yoko Usui, Katsumi Sato, and Shinichi Watabe. Learning hawaiian hula dance by using tablet computer. In *SIGGRAPH Asia 2015 Symposium on Education*, pages 1–2, 2015.

[6] Katerina El Raheb, Marina Stergiou, Akrivi Katifori, and Yannis Ioannidis. Dance interactive learning systems: A study on interaction workflow and teaching approaches. *ACM Comput. Surv.*, 52(3), jun 2019.

[7] Jacob Smith. I can see tomorrow in your dance: A study of dance dance revolution and music video games. *Journal of Popular Music Studies*, 16(1):58–84, 2004.

[8] Jih-Hsuan Lin. "just dance": The effects of exergame feedback and controller use on physical activity and psychological outcomes. *Games for health journal*, 4(3):183–189, 2015.

[9] Jacky CP Chan, Howard Leung, Jeff KT Tang, and Taku Komura. A virtual reality dance training system using motion capture technology. *IEEE transactions on learning technologies*, 4(2):187–195, 2010.

[10] Emiko Charbonneau, Andrew Miller, and Joseph J LaViola Jr. Teach me to dance: exploring player experience and performance in full body dance games. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*, pages 1–8, 2011.

[11] Hironmoy Roy, KUNTALA RAY, Satyajit Saha, and ASIS KUMAR GHOSAL. A study on students' perceptions for online zoom-app based flipped class sessions on anatomy organised during the lockdown period of covid-19 epoch. *Journal of Clinical & Diagnostic Research*, 14(6), 2020.

[12] Michiko Kobayashi. Students'evaluation of google hangouts through a cross-cultural group discussion activity. *Turkish Online Journal of Distance Education*, 16(2):28–39, 2015.

[13] RADiCAL. Artifical intelligence powered 3d motion animation. https://getrad.co/, 2018-2021.

[14] David N Daum and Craig Buschner. Research on teaching blended and online physical education. *Handbook of research on K-12 online and blended learning*, pages 201–222, 2014.

[15] Antonio Camurri, Katerina El Raheb, Oshri Even-Zohar, Yannis Ioannidis, Amalia Markatzi, Jean-Marc Matos, Edwin Morley-Fletcher, Pablo Palacio, Muriel Romero, Augusto Sarti, et al. Wholodance: towards a methodology for selecting motion capture data across different dance learning practice. In *Proceedings of the 3rd International Symposium on Movement and Computing*, pages 1–2.

[16] Craig Buschner. Online physical education: Wires and lights in a box. *Journal of Physical Education, Recreation & Dance*, 77(2):3–8, 2006.

[17] Zihao (Michael) Li. Creativity and opportunity: how covid-19 fosters digital dance education. *Digital Creativity*, 32(3):188–207, 2021.

[18] Roy Martin Simamora. The challenges of online learning during the covid-19 pandemic: An essay analysis of performing arts education students. *Studies in Learning and Teaching*, 1(2):86–103, 2020.

[19] John Hattie and Helen Timperley. The power of feedback. *Review of Educational Research*, 77(1):81–112, 2007.

[20] Richard James, Kerri-Lee Krause, and Claire Jennings. The first year experience in australian universities. *Canberra: Department of Education, Employment and Workplace Relations [DEEWR]*, 2010.

[21] Peter Ferguson. Student perceptions of quality feedback in teacher education. *Assessment & evaluation in higher education*, 36(1):51–62, 2011.

[22] Md Ahea et al. The value and effectiveness of feedback in improving students' learning and professionalizing teaching in higher education. *Journal of Education and Practice*, 7(16):38–41, 2016.

[23] David J Nicol and Debra Macfarlane-Dick. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education*, 31(2):199–218, 2006.

[24] Boud. Enhancing learning through self-assessment. london. routledge. 1995.

[25] Ramli Bakar. The effect of learning motivation on student? s productive competencies in vocational high school, west sumatra. *International Journal of Asian Social Science*, 4(6):722–732, 2014.

[26] Jacob Filgona, John Sakiyo, D Gwany, and Augustine Okoronka. Motivation in learning. *Asian Journal of Education and Social Studies*, 10:16–37, 09 2020.

[27] Boud. Sustainable assessment: rethinking assessment for the learning society. *Studies in continuing education*, 22(2):151–167, 2000.

[28] D Royce Sadler. Formative assessment and the design of instructional systems. *Instructional science*, 18(2):119–144, 1989.

[29] Mantz Yorke. Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher education*, 45(4):477–501, 2003.

[30] Gibbons. Feedback in the dance studio. *Journal of Physical Education, Recreation Dance*, 75:38–43, 09 2004.

[31] Elizabeth Gibbons. Feedback in the dance studio. *Journal of Physical Education, Recreation Dance*, 75:38–43, 09 2004.

[32] Adesola Akinleye and Rose Payne. Transactional space: Feedback, critical thinking, and learning dance technique. *Journal of Dance Education*, 16(4):144–148, 2016.

[33] Andrea E Downie. *Teaching dance: multisensory instruction addresses learning diversity and enhances dance skill learning.* Library and Archives Canada= Bibliotheque et Archives Canada, Ottawa, 2011.

[34] Karen Dearborn and Rachael Ross. Dance learning and the mirror: comparison study of dance phrase learning with and without mirrors. *Journal of Dance Education*, 6(4):109–115, 2006.

[35] Jennifer Goldstein. Easy to dance to: Solving the problems of teacher evaluation with peer assistance and review. *American Journal of Education*, 113(3):479–508, 2007.

[36] Keith Topping. Peer assessment between students in colleges and universities. *Review of educational Research*, 68(3):249–276, 1998.

[37] Yun Xiao and Robert Lucking. The impact of two types of peer assessment on students' performance and satisfaction within a wiki environment. *The Internet and Higher Education*, 11(3-4):186–193, 2008.

[38] Emily Cross, David Kraemer, Antonia Hamilton, William Kelley, and Scott Grafton. Sensitivity of the action observation network to physical and observational learning. *Cerebral cortex (New York, N.Y. : 1991)*, 19:315–26, 06 2008.

[39] Karen Dearborn and Rachael Ross. Dance learning and the mirror: Comparison study of dance phrase learning with and without mirrors. *Journal of Dance Education*, 6(4):109–115, 2006.

[40] Dana Maslovat and Ian Franks. *The Importance of Feedback to Performance*, pages 3–10. 12 2019.

[41] Daniel L Schwartz, Kevin Hartman, et al. It is not television anymore: Designing digital video for learning and assessment. *Video research in the learning sciences*, pages 335–348, 2007.

[42] James Hallam, Emily Keen, Christa Lee, Alison McKenna, and Mudit Gupta. Ballet hero: building a garment for memetic embodiment in dance learning. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers: Adjunct Program*, pages 49–54, 2014.

[43] Slim Essid, Yves Grenier, Mounira Maazaoui, Gaël Richard, and Robin Tournemenne. An audio-driven virtual dance-teaching assistant. In *Proceedings of the 19th ACM international conference on Multimedia*, pages 675–678, 2011.

[44] Julie A Brodie and Elin E Lobel. *Dance and somatics: Mind-body principles of teaching and performance*. McFarland, 2014.

[45] Jacky CP Chan, Howard Leung, Jeff KT Tang, and Taku Komura. A virtual reality dance training system using motion capture technology. *IEEE transactions on learning technologies*, 4(2):187–195, 2010.

[46] Zoe Marquardt, João Beira, Natalia Em, Isabel Paiva, and Sebastian Kox. Super mirror: a kinect interface for ballet dancers. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 1619–1624. 2012.

[47] Jeff KT Tang, Jacky CP Chan, and Howard Leung. Interactive dancing game with real-time recognition of continuous dance moves from 3d human motion capture. In *Proceedings of the 5th international conference on ubiquitous information management and communication*, pages 1–9, 2011.

[48] Liqun Deng, Howard Leung, Naijie Gu, and Yang Yang. Real-time mocap dance recognition for an interactive dancing game. *Computer animation and virtual worlds*, 22(2-3):229–237, 2011.

[49] Yang Yang, Howard Leung, Lihua Yue, and Liqun Deng. Generating a two-phase lesson for guiding beginners to learn basic dance movements. *Computers & Education*, 61:1–20, 2013.

[50] Koaburo Hachimura, Hiromu Kato, and Hideyuki Tamura. A prototype dance training support system with motion capture and mixed reality technologies. In *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)*, pages 217–222. IEEE, 2004.

[51] Andreas Aristidou, Efstathios Stavrakis, Panayiotis Charalambous, Yiorgos Chrysanthou, and Stephania Loizidou Himona. Folk dance evaluation using laban movement analysis. *Journal on Computing and Cultural Heritage (JOCCH)*, 8(4):1–19, 2015.

[52] Sarah Fdili Alaoui, Kristin Carlson, and Thecla Schiphorst. Choreography as mediated through compositional tools for movement: Constructing a historical perspective. In *Proceedings of the 2014 International Workshop on Movement and Computing*, pages 1–6, 2014.

[53] ROSANNE SPECTOR. Dancing with data adds to the show, 2015. Last accessed 10 December 2021.

[54] Akio Nakamura, Sou Tabata, Tomoya Ueda, Shinichiro Kiyofuji, and Yoshinori Kuno. Dance training system with active vibro-devices and a mobile image display. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3075–3080. IEEE, 2005.

[55] Nadia Magnenat-Thalmann, Dimitrios Protopsaltou, and Evangelia Kavakli. Learning how to dance using a web 3d platform. In *International Conference on Web-Based Learning*, pages 1–12. Springer, 2007.

[56] Guoyu Sun, Paisarn Muneesawang, M Kyan, Haiyan Li, Ling Zhong, Nan Dong, Bruce Elder, and Ling Guan. An advanced computational intelligence system for training of ballet dance in a cave virtual reality environment. In *2014 IEEE International Symposium on Multimedia*, pages 159–166. IEEE, 2014.

[57] Doo Young Kwon and Markus Gross. Combining body sensors and visual sensors for motion training. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pages 94–101, 2005.

[58] Willy Pieter. Taekwondo. In *Combat sports medicine*, pages 263–286. Springer, 2009.

[59] Vicky Karkou, Sophia Bakogianni, and Evangelia Kavakli. Traditional dance, pedagogy and technology: an overview of the webdance project. *Research in Dance Education*, 9(2):163–186, 2008.

[60] Dimitrios S Alexiadis, Philip Kelly, Petros Daras, Noel E O'Connor, Tamy Boubekeur, and Maher Ben Moussa. Evaluating a dancer's performance using kinect-based skeleton tracking. In *Proceedings of the 19th ACM international conference on Multimedia*, pages 659–662, 2011.

[61] Fraser Anderson, Tovi Grossman, Justin Matejka, and George Fitzmaurice. Youmove: enhancing movement training with an augmented reality mirror. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pages 311–320, 2013.

[62] Alberto Alvisi, Alessandro Narduzzo, and Marco Zamarian. Playstation and the power of unexpected consequences. *Information, Communication Society*, 6:608–627, 12 2003.

[63] M Gregor, Petr Horejsi, and M Simon. Case study: Motion capture for ergonomics. In *Proceedings of the 25th International Business Information Management Association Conference–Innovation Vision 2020: From Regional Development Sustainability to Global Economic Growth, IBIMA 2015*, pages 468–476, 2015.

[64] Ronald Sidharta and Carolina Cruz-Neira. Cyclone uppercut, a boxing game for an immersive environment. pages 363–364, 01 2005.

[65] Christos Mousas. Performance-driven dance motion control of a virtual partner character. pages 57–64, 03 2018.

[66] Guido Makransky, Stefan Borre-Gude, and Richard E Mayer. Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *Journal of Computer Assisted Learning*, 35(6):691–707, 2019.

[67] Göran Fransson, Jörgen Holmberg, and Claes Westelius. The challenges of using head mounted virtual reality in k-12 schools from a teacher perspective. *Education and Information Technologies*, 25(4):3383–3404, 2020.

[68] Lisa Rebenitsch and Charles Owen. Review on cybersickness in applications and visual displays. *Virtual Reality*, 20(2):101–125, 2016.

[69] Antonio Camurri, Corrado Canepa, Nicola Ferrari, Maurizio Mancini, Radoslaw Niewiadomski, Stefano Piana, Gualtiero Volpe, Jean-Marc Matos, Pablo Palacio, and Muriel Romero. A system to support the learning of movement qualities in dance: a case study on

dynamic symmetry. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 973–976, 2016.

[70] Sarah Fdili Alaoui, Frédéric Bevilacqua, Bertha Bermudez Pascual, and Christian Jacquemin. Dance interaction with physical model visuals based on movement qualities. *International Journal of Arts and Technology*, 6(4):357–387, 2013.

[71] Stefano Piana, Paolo Alborno, Radoslaw Niewiadomski, Maurizio Mancini, Gualtiero Volpe, and Antonio Camurri. Movement fluidity analysis based on performance and perception. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 1629–1636, 2016.

[72] Rudolf Laban and Lisa Ullmann. The mastery of movement. 1971.

[73] Mitsu Yoshimura, Norio Mine, Tamiko Kai, and I Yoshimura. Quantification of characteristic features of japanese dance for individuality recognition. In *Proceedings 10th IEEE International Workshop on Robot and Human Interactive Communication. ROMAN 2001 (Cat. No. 01TH8591)*, pages 188–193. IEEE, 2001.

[74] Gang Qian, Feng Guo, Todd Ingalls, Loren Olson, Jodi James, and Thanassis Rikakis. A gesture-driven multimodal interactive dance system. In *2004 IEEE International Conference on Multimedia and Expo (ICME)(IEEE Cat. No. 04TH8763)*, volume 3, pages 1579–1582. IEEE, 2004.

[75] Dieter Drobny, Malte Weiss, and Jan Borchers. Saltate! a sensor-based system to support dance beginners. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, pages 3943–3948. 2009.

[76] Alexandros Kitsikidis, Kosmas Dimitropoulos, Erdal Yilmaz, Stella Douka, and Grammalidis Nikos. Multi-sensor technology and fuzzy logic for dancer's motion analysis and performance evaluation within a 3d virtual environment. 06 2014.

[77] Gayle Kassing. *History of Dance, 2E.* Human Kinetics, 2017.

[78] Rachel Evelyn Ward. Biomechanical perspectives on classical ballet technique and implications for teaching practice. *University of New South Wales*, 2012.

[79] Iyengar Mukunda Kalpana. Bharatanatyam and mathematics: Teaching geometry through dance. *Journal of Fine and Studio Art*, 5(2):6–17, 2015.

[80] Carrie Pledger. Ballroom dance: An education like no other. *Inquiry*, 20(1):61–74, 2016.

[81] Esha Bhattacharya Chatterjee. Contemporary bollywood dance: Analyzing it through the interplay of social forces. *International Journal of Scientific and Research Publications*, 6(6):54–60, 2016.

[82] Tamara Munzner. *Visualization analysis and design*. CRC press, 2014.

[83] Eric N Franklin. *Dance imagery for technique and performance*. Human Kinetics, 2013.

[84] M Guadagnoli, W Holcomb, and M Davis. The efficacy of video feedback for learning the golf swing. *Journal of sports sciences*, 20(8):615–622, 2002.

[85] Heather Kelley and Raymond G Miltenberger. Using video feedback to improve horseback-riding skills. *Journal of Applied Behavior Analysis*, 49(1):138–147, 2016.

[86] Caryl H Hitchcock, Peter W Dowrick, and Mary Anne Prater. Video self-modeling intervention in school-based settings: A review. *Remedial and Special Education*, 24(1):36–45, 2003.

[87] Mary Anne Prater, Nari Carter, Caryl Hitchcock, and Peter Dowrick. Video self-modeling to improve academic performance: A literature review. *Psychology in the Schools*, 49(1):71–81, 2012.

[88] Peter W Dowrick. A review of self modeling and related interventions. *Applied and preventive psychology*, 8(1):23–39, 1999.

[89] Scott Bellini and Jennifer Akullian. A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Exceptional children*, 73(3):264–287, 2007.

[90] Richard A Schmidt. Frequent augmented feedback can degrade learning: Evidence and interpretations. In *Tutorials in motor neuroscience*, pages 59–75. Springer, 1991.

[91] John Annett. Feedback and human behaviour: The effects of knowledge of results, incentives and reinforcement on learning and performance. 1969.

[92] Lawrence Karlin and Rudolf G Mortimer. Effect of verbal, visual, and auditory augmenting cues on learning a complex motor skill. *Journal of experimental psychology*, 65(1):75, 1963.

[93] J Patrick and F Mutlusoy. The relationship between types of feedback, gain of a display and feedback precision in acquisition of a simple motor task. *The Quarterly Journal of Experimental Psychology Section A*, 34(1):171–182, 1982.

[94] Wanmo Koo. Usage of smartphone applications: A descriptive study of top 100 us retailers. *International Journal of Interactive Mobile Technologies*, 10(3), 2016.

[95] Maddock Meredith, Steve Maddock, et al. Motion capture file formats explained. *Department of Computer Science, University of Sheffield*, 211:241–244, 2001.

[96] Tuukka Takala, Meeri Mäkäräinen, and Perttu Hamalainen. Immersive 3d modeling with blender and off-the-shelf hardware. pages 191–192, 03 2013.

[97] Leo Liberti, Carlile Lavor, Nelson Maculan, and Antonio Mucherino. Euclidean distance geometry and applications. *SIAM Review*, 56, 05 2012.

[98] Graphics and University Of Cyprus Virtual Reality Lab. Dance motion capture database, 2015. Last accessed 10 December 2021.

[99] Leyde Briceno and Gunther Paul. Makehuman: a review of the modelling framework. In *Congress of the International Ergonomics Association*, pages 224–232. Springer, 2018.

[100] Zhaohui Cen, Jiaolong Wei, and Rui Jiang. A grey-box neural network based identification model for nonlinear dynamic systems. In *The Fourth International Workshop on Advanced Computational Intelligence*, pages 300–307. IEEE, 2011.

[101] Simone Wirtz, Eva-Maria Jakobs, and Martina Ziefle. Age-specific usability issues of software interfaces. In *Proceedings of the IEA*, volume 17, 2009.

[102] Ravi Chandra Chaitanya Guntupalli. User interface design: methods and qualities of a good user interface design, 2008.

[103] Theo Mandel. *The elements of user interface design*, volume 20. Wiley New York, 1997.

[104] Sommerville Ian. Software engineering eighth edition. *Harlow: Pearson Education Limited*, 2007.

[105] Clayton Lewis, Peter Polson, and Tim McKay. *Human-Computer Interface Design: success stories, emerging methods, and real-world context*. Morgan Kaufmann Pub, 1996.

[106] Brenda Laurel. The art of human-computer interface design. 1990.

[107] Rolf Molich and Jakob Nielsen. Improving a human-computer dialogue. *Communications of the ACM*, 33(3):338–348, 1990.

[108] Hendrik Müller, Aaron Sedley, and Elizabeth Ferrall-Nunge. Survey research in hci. In *Ways of Knowing in HCI*, pages 229–266. Springer, 2014.

[109] A Ant Ozok. Survey design and implementation in hci. In *The human-computer interaction handbook*, pages 1177–1196. CRC Press, 2007.

[110] Mick P Couper. Web surveys: A review of issues and approaches. *The Public Opinion Quarterly*, 64(4):464–494, 2000.

Appendix A

BVH (BioVision hierarchial data)file

Motion capture systems capture live motions by tracking a series of key locations in space over time and converting them into a three-dimensional digital representation. The captured subject can be anything that exists in the actual world, with key points placed on the item to best reflect the orientations of the object's moving pieces, such as joints or pivot points [95]. Motion Capture data has been used in television commercials and computer games. There are multiple formats for motion capture data. BioVision Hierarchial data (BVH) is one of them.Biovision, a motion capture services provider, created the BVH file format as a mechanism to send motion capture data to its customers. Biovision hierarchical data is referred to as BVH. This format largely superseded a previous format they created, the Biovision Action data format, as a mechanism to convey skeletal hierarchy information in addition to motion data.The skeletal structure of the Biovision Hierarchial data file can be seen in Figure 4.2.

The following terminology has been used to describe the motion capture data and processing:

a) Skeleton - The entire persona who is represented by the motion.

b) Bone - The most fundamental component of a skeleton. During the animation, each bone represents the smallest portion of the motion that is sensitive to individual translation and orientation modifications. A skeleton is made up of several bones (typically in a hierarchical structure, as shown in Figure 4.2), each of which may be linked to a vertex mesh to represent a specific section of the character, such as the femur or humerus.

b) Hierarchy - A BVH file is divided into two sections: a header section that explains the skeleton's hierarchy and beginning position, and a data portion that provides the motion data.The keyword ”HIERARCHY” appears at the start of the header section. The keyword ”HIERARCHY” appears at the start of the header section. A BVH file can theoretically include any number of hierarchies. The format of the motion section restricts the number of segments in practise; one sample in time for all segments is on one line of data [95].

c) Root - Hierarchy is followed by the ”Root”.The ROOT keyword indicates the start of a new skeletal hierarchical structure.The ”Root” bone is the initial start, and all the translation and rotation of different bones are compared involving the root bone.

d) Channel or Degree of Freedom (DOF) - Over the course of the animation, each bone in a skeleton can have its location, orientation, scale change, with each parameter being referred to as a channel (or DOF). The animation is created by changes in the channel data over time.

e) Frame - Every animation is made up of a number of frames, each of which has its own set of channel data for each bone. Motion capture data can be acquired at speeds of up to 240 frames per second, but in most cases, 30 or 60 frames per second is the usual.Motions with high frequency material, such as a combination of karate movements, are captured at high frame rates. Although the extra detail is often not visible during real-time playback because to the maximum refresh rates of display hardware, it may be beneficial for adding motion blurring to the animation or just for motion analysis.

f) Frame Time - The ”Frame Time” definition appears after the frames definition and indicates the data sampling rate. If the sample rate in the BVH file is 0.033333, the frame time is 30 frames per second, which is the typical rate of sampling in a BVH file.

g) Offset - The offset is specified by the keyword ”OFFSET” and it indicates the X,Y and Z offset of the segment from its parent. The offset information also indicates the length and direction used for drawing the parent segment. The first piece of information of a segment is the offset of that segment from its parent, or in the case of the root object the offset will generally be zero.

h) Joint - A joint definition is almost similar to the root definition except for the number of channels. This is where the recursion takes place, the rest of the parsing of the joint information proceeds just like a root. The recursion occurs here, and the remainder of the parsing of the joint information proceeds as if it were a root.

i) Motion - The motion section begins with a single line containing the keyword ”MOTION.”This is followed by a line indicating the number of frames in the file, which includes the ”Frames” and ”Frame Time”. The motion data is contained in the rest of the file. Each line represents a single motion data sample. As the skeleton hierarchy is parsed, the numbers appear in the order of the channel requirements.

A.1 Interpreting a BVH file

```
HIERARCHY
ROOT Hips
{
    OFFSET 0.000000 96.750603 -1.056153
    CHANNELS 6 Xposition Yposition Zposition Xrotation Yrotation Zrotation
    JOINT RightUpLeg
    {
        OFFSET -9.005805 -1.450562 0.526166
        CHANNELS 6 Xposition Yposition Zposition Xrotation Yrotation Zrotation
        JOINT RightLeg
        {
            OFFSET -4.555083 -42.308475 -1.278553
            CHANNELS 3 Xrotation Yrotation Zrotation
            JOINT RightFoot
            {
                OFFSET -3.507217 -39.548161 -6.278737
                CHANNELS 6 Xposition Yposition Zposition Xrotation Yrotation Zrotation
                JOINT RightToeBase
                {
                    OFFSET 0.070749 -10.430739 16.592361
                    CHANNELS 6 Xposition Yposition Zposition Xrotation Yrotation Zrotation
                    JOINT RightToeBase_End
                    {
                        OFFSET 0.058216 0.008661 6.234280
                        CHANNELS 3 Xrotation Yrotation Zrotation
                        End Site
                        {
                            OFFSET 0.058216 0.008661 6.234282
                        }
                    }
                }
            }
        }
    }
```

Figure A.1: BVH file interpretation

A clear picture of the BVH file format can be seen in Figure A.1. The file's hierarchical section begins with the term HIERARCHY, which is followed by the keyword ROOT and the name of the bone that is the skeletal hierarchy's root on the next line. The ROOT keyword signals the beginning of a new skeletal hierarchical structure, and while the BVH file can include several skeletons, it is more common to have only one skeleton declared per file.

The skeleton's remaining structure is specified in a recursive manner, with the definition of each BONE,including any children, contained in curly braces, which are delimited on the preceding line by the keyword JOINT (or ROOT in the case of the root bone), followed by the name of the bone. It's a good idea to indent the bone's content (with a tab) and match the closing curly brace with the equivalent opening one when introducing a left curly brace.Furthermore, while the hierarchical indentation isn't required, it does help to make the file more legible for humans.

The first line of each bone's definition, bounded by the term OFFSET, describes the translation of the bone's origin along the x, y, and z-axes with respect to its parent's origin (or globally in the case of the root bone). The offset also serves the goal of implicitly specifying the length and direction of the parent's bone; however, establishing the length and direction of a bone with several offspring presents a dilemma.

The term CHANNELS is prefixed to the second line of a bone's definition, which defines the current bone's DEGREE OF FREEDOMs. The sequence in which the channels are displayed is crucial in two ways. For starters, the order in which each channel appears in the hierarchy portion of the file corresponds to the order in which the data appears in the motion section of the file(explained in Figure A.2).

The motion part of the file (Figure A.2), for example, provides information for the channels of the root bone in the sequence defined in the hierarchy,[95] followed by channel data for its first child, followed by channel data for that child, and so on.The second thing to remember about channel ordering is that when building the bone's rotation matrix, the concatenation order of the Euler angles must match the order shown in the CHANNEL section. As the Euler order is set for each bone, it is possible to have various orders for different bones, which must be taken into account in order to get a correct-looking animation. In some cases, rotational channels are specified differently for different joints.The lines following the bone specification, starting with the keyword JOINT, are used to specify child items, however in the case of end-effectors, a special tag is used, ”End Site,” which contains an OFFSET triple that is used to infer the bone’s length and orientation.

```
                    End Site
                    {
                        OFFSET -0.058216 0.008661 6.234205
                    }
                }
            }
        }
    }
}
MOTION
Frames: 400
Frame Time: 0.033333
0.000000 96.750603 -1.056153 -0.000000 0.000000 -0.000000 -9.005805 -1.450546 0.526165 -0.000000 -0.000001 -0.000000 -0.00000
0.160564 94.521843 -0.620962 6.813508 0.202267 -1.383781 -9.005805 -1.450550 0.526161 -12.802428 -8.397829 5.771698 19.856040
0.163017 94.525864 -0.619972 6.824173 0.342544 -1.380282 -9.005805 -1.450570 0.526161 -12.821079 -8.400127 5.743317 19.853533
0.164924 94.532730 -0.620193 6.834104 0.479377 -1.368395 -9.005808 -1.450550 0.526162 -12.843256 -8.404628 5.703495 19.849594
0.167145 94.542755 -0.621013 6.840520 0.613387 -1.347545 -9.005807 -1.450550 0.526164 -12.864940 -8.407866 5.654181 19.846088
0.167385 94.555046 -0.622722 6.838516 0.600697 -1.304766 -9.005808 -1.450548 0.526163 -12.877952 -8.406587 5.603035 19.842212
0.166249 94.568771 -0.621611 6.829237 0.591594 -1.260679 -9.005809 -1.450545 0.526163 -12.882582 -8.403481 5.556142 19.839017
0.163917 94.580650 -0.616101 6.815077 0.587789 -1.217084 -9.005808 -1.450538 0.526163 -12.882009 -8.398682 5.515523 19.838349
0.159628 94.588486 -0.605111 6.798004 0.590539 -1.175753 -9.005807 -1.450549 0.526161 -12.881321 -8.394368 5.480259 19.847191
0.152688 94.590591 -0.591467 6.777480 0.598034 -1.138558 -9.005808 -1.450547 0.526162 -12.880564 -8.390706 5.450402 19.867297
0.142740 94.586060 -0.577277 6.750489 0.612957 -1.105630 -9.005808 -1.450559 0.526162 -12.876295 -8.383633 5.424251 19.902661
```

Figure A.2: BVH file showing MOTION

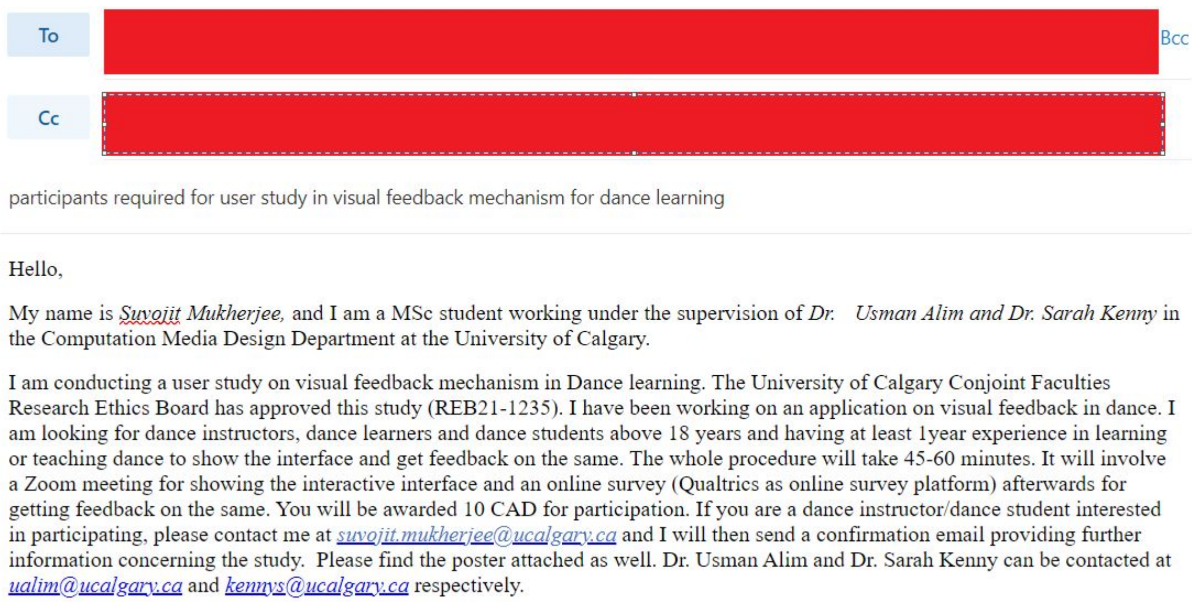
The number of frames in the animation, frame rate, and channel data are all contained in the second part of a BVH file, which is designated by the term MOTION.The line holding the number of frames begins with the keyword ”Frames” and is followed by the number of frames as a positive decimal integer (rather than hexadecimal or octadecimal). The frame rate is displayed on a line that begins with ”Frame Time” and ends with a positive float indicating the length of a single frame. Each line of float values represents one animation frame in this definition.

Appendix B

Survey Details

B.1 Survey posters

Recruitment posters (figure B.2) and emails (figure B.1) were sent mentioning the selection criteria for participant selection.



To [redacted] Bcc

Cc [redacted]

participants required for user study in visual feedback mechanism for dance learning

Hello,

My name is *Suvojit Mukherjee,* and I am a MSc student working under the supervision of *Dr. Usman Alim and Dr. Sarah Kenny* in the Computation Media Design Department at the University of Calgary.

I am conducting a user study on visual feedback mechanism in Dance learning. The University of Calgary Conjoint Faculties Research Ethics Board has approved this study (REB21-1235). I have been working on an application on visual feedback in dance. I am looking for dance instructors, dance learners and dance students above 18 years and having at least 1year experience in learning or teaching dance to show the interface and get feedback on the same. The whole procedure will take 45-60 minutes. It will involve a Zoom meeting for showing the interactive interface and an online survey (Qualtrics as online survey platform) afterwards for getting feedback on the same. You will be awarded 10 CAD for participation. If you are a dance instructor/dance student interested in participating, please contact me at *suvojit.mukherjee@ucalgary.ca* and I will then send a confirmation email providing further information concerning the study. Please find the poster attached as well. Dr. Usman Alim and Dr. Sarah Kenny can be contacted at *ualim@ucalgary.ca* and *kennys@ucalgary.ca* respectively.

Figure B.1: Recruitment email

Figure B.2: Recruitment poster

B.2 Survey questions

Its important to carefully think and design survey so that it aligns with the research goal. Poor questionnaire design may introduce measurement error , defined as the deviation of the respondents' answers from their true values on the measure [108]. According to Couper and Mick [110], measurement error in self-administered surveys can arise from the respondent (e.g., lack of motivation, comprehension problems, deliberate distortion) or from the instrument (e.g., poor wording or design, technical flaws). In most surveys, there is only one option to deploy, and unlike qualitative research, there is no possibility for explanation or inquiry. As a result, it's critical that the questions appropriately match the research objective.

B.2.1 Survey questions for 'Danceshala'

I designed the survey questions keeping in mind our research goal which is to find out whether our interactive interface will actually help beginners for learning dance online. Every element in the feedback interface has been well-thought-out and we needed to know how the dancers or dance-teachers feel about them. Keeping all this in mind, the questions framed for the survey are:

1. 1st question (Age-range) : In the first question, we wanted to know about the participant's age range (Figure B.3). The survey is totally anonymous and we wanted to know how people of different age range react to the same.

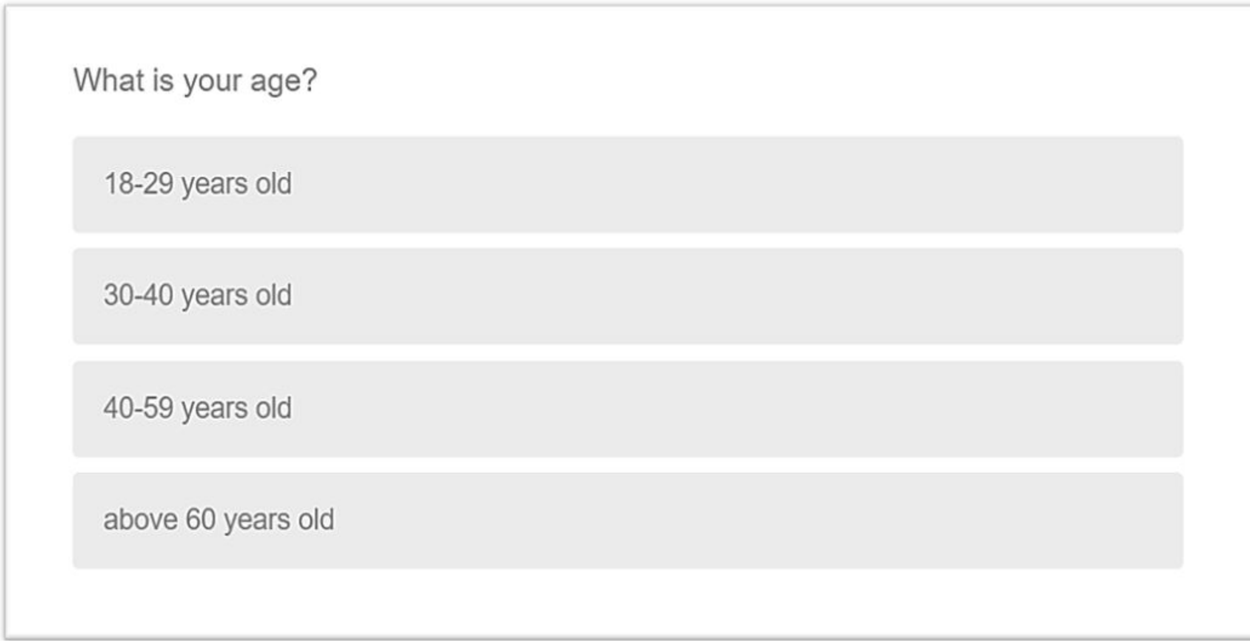


Figure B.3: 1st question - age-range

2. 2nd question (Experience in dance) : In the second question, I wanted to find out the dance experience of the participant (Figure B.5). I wanted to cover majority of the population and therefore, I put up three options from which the respondent can select only one.

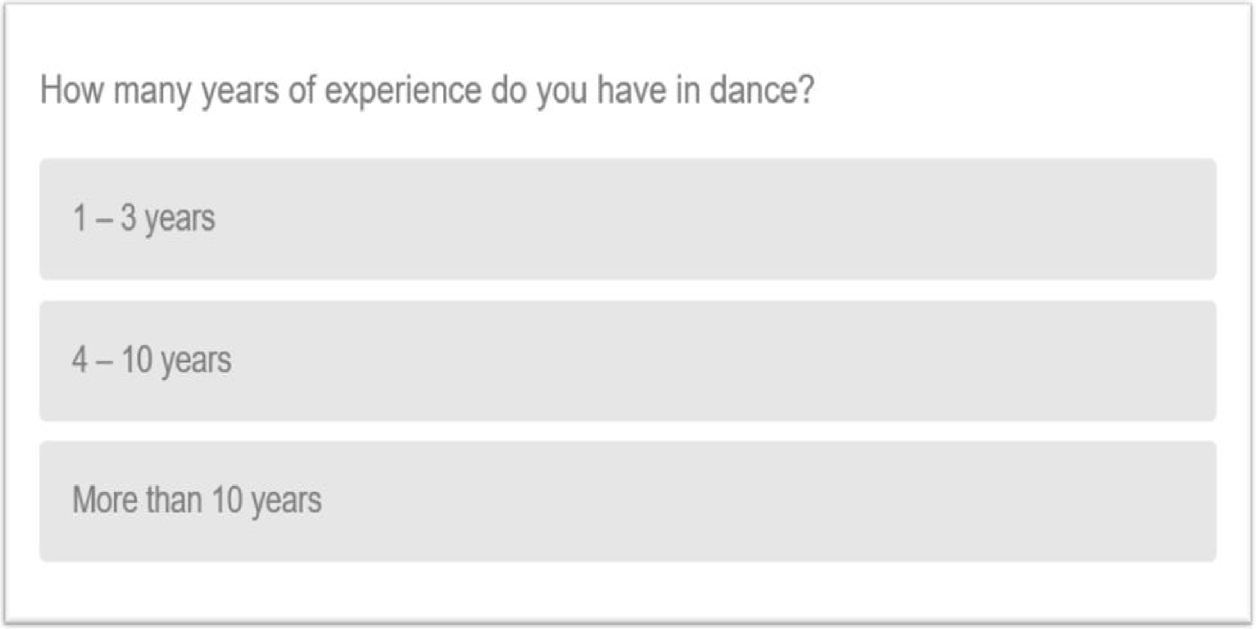


Figure B.4: 2nd question - Experience in dance

3. 3rd question (how easy is it for a beginner to understand the information displayed) : This interface has been designed to assist the beginners in dance. Therefore, it is imperative for the researcher to understand if users will be able to understand the information displayed on the interface.

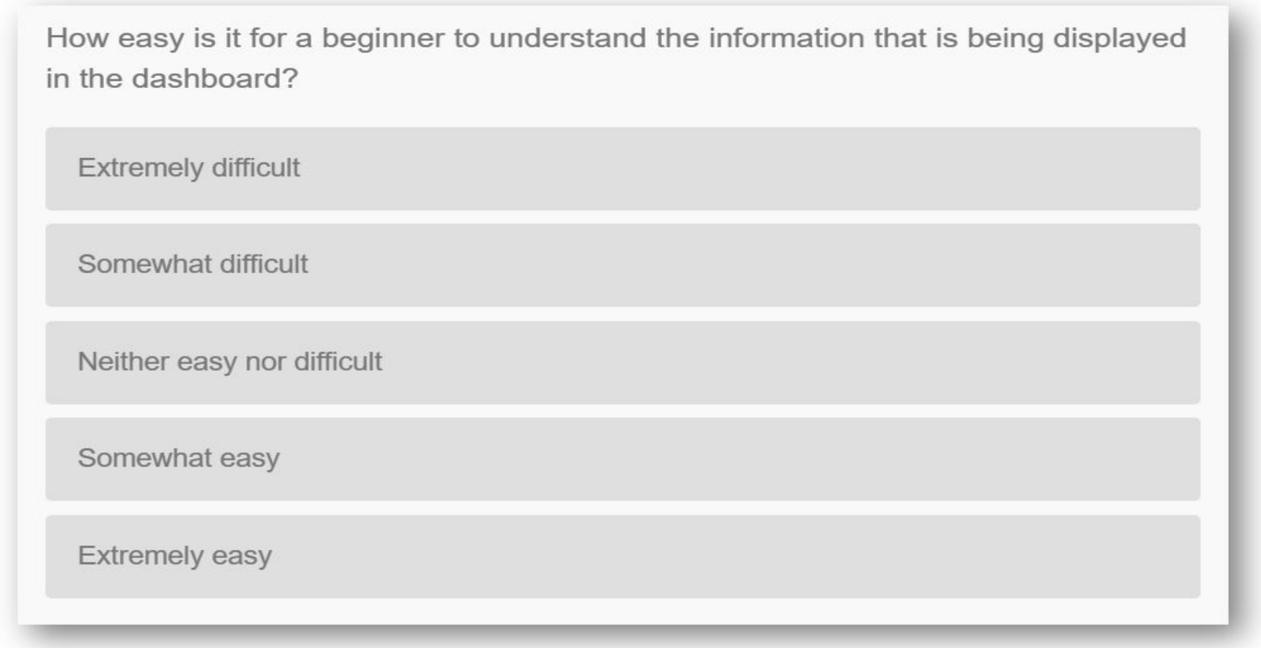


Figure B.5: 3rd question - how easy is it for a beginner to understand the information displayed

4. 4th question (Whether the visual feedback provided by the interface will really help a beginner to learn dance) : This is a ranking question (Figure B.6) where the participant needs to rank the effectiveness of the visual feedback interface in teaching dance to beginners. This is where the dancers and dance teachers feedback can really assist to gauge the potency of the interface. The dance-teachers teach beginners regularly and the dancers have also crossed the initial stage of being a beginner. Hence, they can use their experience to correctly identify the efficacy of the interface. The user can only select one option among the five options.

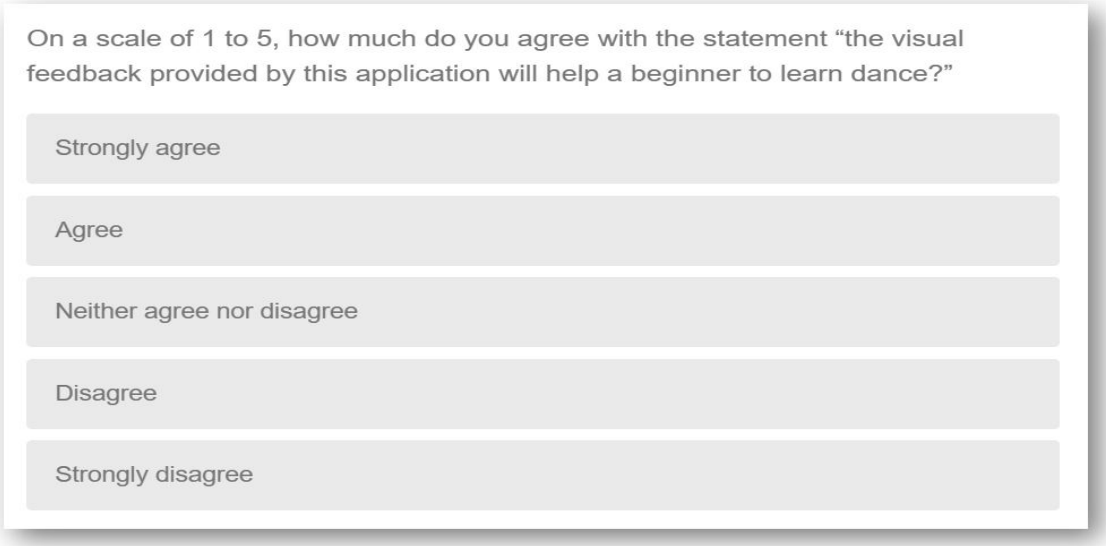


Figure B.6: 4th question - How effective is the visual feedback provided in teaching dance to a beginner

5. 5th question(How easy it is for a beginner to control the different interface elements) : The interface has various elements like the score, the analysis, the user and expert videos and video controls, main graph and accuracy graph. We wanted to know if a beginner really understood the controls and the interactive features. For this reason, I created this close-ended question with 5 options(Figure B.7).

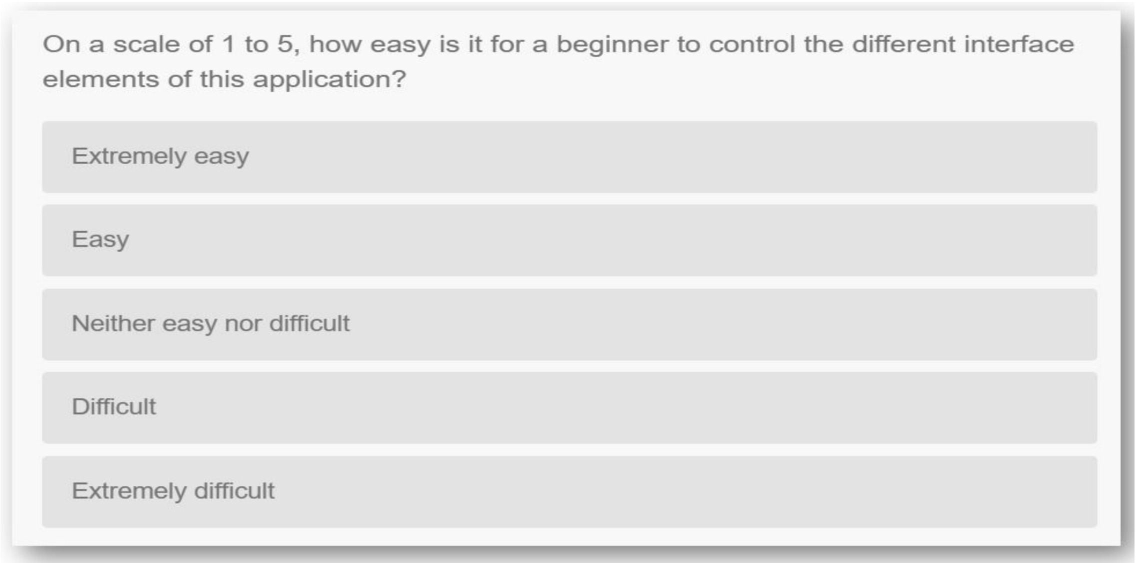


Figure B.7: 5th question - How easy it is for a beginner to control the different interface elements

6. 6th question(How useful does the participant find the level of interaction provided by different elements of the application?) : Interactivity is one of the most important factors of our interface design. Thus, I framed this question (Figure B.8) to find whether participant finds the level of interaction equally interesting and useful. It is also a close-ended question with 5 options.

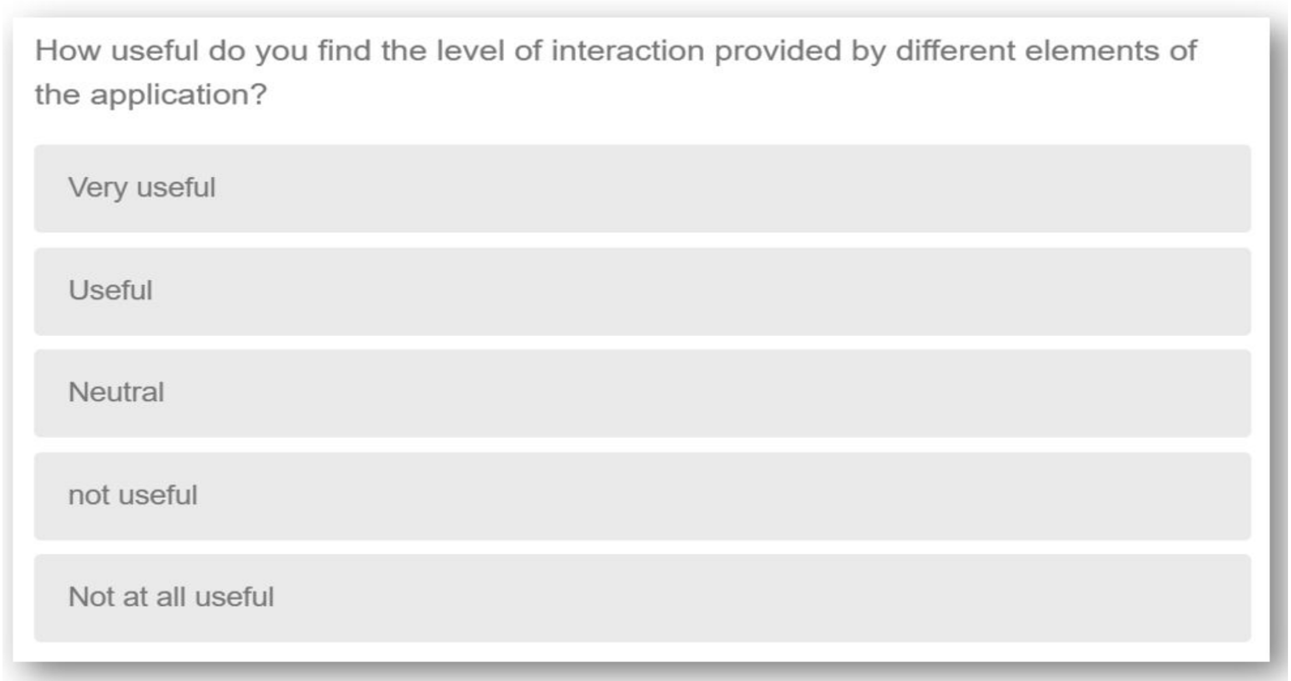


Figure B.8: 6th question - (How useful does the participant find the level of interaction provided by different elements of the application?)

7. 7th question(Which visual element do you think is the most beneficial for a beginner?) : The interface has various elements like Main graph, Error bar chart or accuracy graph, Video and Video controls and score and analysis.With the help of the seventh question (Figure B.9), we wanted to know which visual element is the most advantageous for dance-learning.

Figure B.9: 7th question - (Which visual element do you think is the most beneficial for a beginner?)

8. 8th question(Please let us know if you have any suggestions for improvement.) : The 8th question (Figure B.10 asks for the user's feedback and improvement suggestions. This is an open-ended question.

Figure B.10: 8th question - (User feedback suggestions)