DESIGN GUIDELINES FOR THE DEVELOPMENT OF MOTION CAPTURE SYSTEM FOR ELEMENTARY SCHOOL CLASSROOM INTEGRATION

A Thesis

by

KUMAR SRIDHARAMURTHY

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Francis Quek
Co-Chair of Committee,	Sharon Lynn Chu
Committee Member,	John Keyser
Head of Department,	Tim McLaughlin

December 2015

Major Subject: Visualization

ABSTRACT

Motion capture technology has revolutionized entertainment and gaming industries. Research has shown that the motion capture technology also has the potential to impact education and help kinesthetic learners. The goal of this thesis is to come up with the design guidelines for developing such a motion capture system for elementary school classroom integration.

An exemplar system called the Digital Micro-Enactment (DiME) marker based system was used to study the feasibility of motion capture system in a classroom setting. A focus group was conducted with 4 elementary school teachers to understand the constraints of a classroom. The discussion was analyzed to formulate the design guidelines for the development of DiME markerless motion capture system. This system was compared against DiME marker based system in a user study with 6 elementary school teachers. Quantitative and qualitative data analysis indicated that DiME markerless system was preferred by the teachers over DiME marker based system. This thesis will benefit educators and researchers by providing the design guidelines for developing motion capture systems inside classrooms.

DEDICATION

This thesis is dedicated to my father, A. V. Sridharamurthy, who taught me to dream big and to be fearless, which has brought me from a remote town in India to where I am today.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Francis Quek, and committee members, Dr. Sharon Lynn Chu and Dr. John Keyser, for their support and for helping me understand the difference between application development and research.

I would also like to thank my mother Poornima Murthy and sisters Akhila Sridharamurthy, Aparna Sridharamurthy and Smitha Sridharamurthy for their support and care. Thanks to my wonderful friend, Michael Saenz for his valuable feedback during crunch time. Thanks to the department faculty and staff for making my time at Texas A&M University a thrilling experience. I want to thank Lavanya Rao for the patient listening and moral support. I also want to extend my gratitude to all the elementary school teachers who participated in the user study.

NOMENCLATURE

DiME	Digital Micro Enactment System
Mocap	Motion Capture
PS	PlayStation
ELA	English Language Arts
B/CS	Bryan College Station
PC	Personal Computer
SUS	System Usability Scale
STAAR	State of Texas Assessment of Academic Readiness

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iii

ACKNOWLEDGEMENTS iv

NOMENCLATURE...v

TABLE OF CONTENTS vi

LIST OF FIGURES.. viii

LIST OF TABLES ix

1. INTRODUCTION..1

1.1. Motivation 2
1.2. Aim of the work 3

2. RELATED WORK..4

2.1. Technology integration in classrooms...4
2.2. Commodity motion capture systems 5
2.3. Comparison of marker-based and markerless motion capture systems 8
2.4. Motion capture systems in education 10
2.5. Digital Micro Enactment (DiME) marker based system..............................11

3. METHODOLOGY 15

3.1. Formative study...15
3.2. Findings of the study 16
3.3. Design criteria 22

4. IMPLEMENTATION 24

4.1. DiME markerless system 24
4.2. Choice of hardware 25
4.3. Architecture overview 28
4.4. Calibration technique 29

4.5. Comparison between the DiME systems 32

5. EVALUATION 33

5.1. Research questions 33
5.2. Study design 33
5.3. Quantitative feedback 36
5.4. Lesson plan 38
5.5. Qualitative feedback 39
5.6. Design guidelines 45

6. DISCUSSION 46

6.1. Significance of the design guidelines for developers 46
6.2. Classroom integration 46

7. CONCLUSION AND FUTURE WORK 48

7.1. Conclusion 48
7.2. Future work 48

REFERENCES 50

APPENDIX 56

LIST OF FIGURES

Page

Figure 1 : Optitrack motion capture setup with Flex 13 cameras 6

Figure 2 : Microsoft Kinect for Windows v2 sensor..8

Figure 3: Body gestures made by the actor holding a physical object (Left), animated character with a virtual prop (right)..13

Figure 4: Motion capture input for DiME marker-based and DiME markerless 24

Figure 5: DiME markerless system hardware 25

Figure 6: Sony PlayStation Move controller..27

Figure 7: DiME markerless system hardware setup...28

Figure 8: Samples taken with the PS Move controller on the right hand........................30

Figure 9: SUS scores for DiME marker based and DiME markerless for 6 participants 37

Figure 10: Teachers' response to whether they thought it was practical to use DiME system 40

LIST OF TABLES

Page

Table 1: Teacher demographics for the formative study...17

Table 2: Design criteria for motion capture systems...22

Table 3: Comparison of the design criteria for motion capture systems..........................32

Table 4: Teacher demographics for the evaluative user study 35

Table 5: Statistically significant statements from SUS 38

Table 6: Design guidelines along with the frequency of repetitions................................41

Table 7: Final design guidelines...45

1. INTRODUCTION

The advent of technology has made it possible to equip classrooms today with computers and tablets. While touch screen technologies are getting introduced through SMART boards inside classrooms, other technologies like motion sensing are expected to impact classroom education in the next three years [1]. Advancement in motion sensing input devices has made it possible to incorporate motion capture based technologies in schools. It has been shown that motion based technology can improve children's sense of self efficacy in storytelling [2]. It has also been used to improve the quality of Chemistry, Biology and Physical education in schools [3, 4]. It can accommodate multi-user learning experiences which is based on theories of embodied learning. There has been a lot of research about plausible benefits of using motion capture systems in classroom teaching but many of these studies also point to the pedagogical and technical constraints associated with implementing this technology [5]. This research aims to address some of those constraints and provide design guidelines for integrating motion capture technology inside classroom. These design guidelines are tested by developing an exemplar system called the Digital Micro Enactment (DiME) markerless system. An evaluative study was conducted to test this exemplar system with 6 school teachers using System Usability Scale (SUS) and the results are discussed. 9 design guidelines emerged out of this study which would help developers and researchers to design systems that use motion capture technology in classroom environment.

1.1. Motivation

Motion capture is the process of recording movements of objects or people. This technique is extensively used in the entertainment industry where the performance of an actor is recorded and applied to a 3D model. Motion capture technology has evolved over the years and is being used in military, sports, medical and educational applications. The introduction of commodity input devices like the Nintendo Wii, PlayStation Move and Microsoft Kinect has revolutionized the use of motion capture technology [6] [7]. The devices that were mainly built for motion based gaming have been used by the teachers to accomplish educational objectives in Schools. In [8] children found Kinect interaction more engaging than using traditional mouse and keyboard approach. Report on key trends in educational technology indicates that there will be an emphasis on intuitive, interactive learning experiences in the next five years. Also, touch screen and gesture-based technologies are expected to play a more significant role in creating innovative technology-rich learning environments [1].

It is high time that we leverage the advancements in technology to augment classroom education in schools. The goal of this research is to investigate different motion capture systems and provide the design guidelines to develop a motion capture system inside classroom environment. This also poses various design and technical challenges that will be addressed. Solving these challenges will result in a motion capture system that is scalable, portable and one that does not need technical expertise to implement in the classroom.

1.2. Aim of the work

The purpose of this research is to inform the developer community about designing motion capture systems in schools. The following three steps were taken to accomplish this goal. Firstly, some design constraints of using motion capture technology inside classrooms were observed by interviewing elementary school teachers. Secondly, a list of design criteria that must be taken into account while designing motion capture based systems were formulated. Thirdly, the design criteria was evaluated by conducting a usability study with elementary school teachers.

2. RELATED WORK

2.1. Technology integration in classrooms

Technology integration in classroom is an established area of research in which there are many frameworks for technology integration. Groff. J et al, [9] introduced i5 (Individualized Inventory for Integrating Instructional Innovations), to help teachers predict the chances of success of technology integration. It takes into account the nature of the schools, teachers, students and projects to determine a rating that predicts the likelihood of successful technology integration. Wang. Q. proposed another framework [10] to guide the integration of Information and Communication Technology (ICT) into teaching and learning based on three fundamental elements: pedagogy, social interaction and technology. Inan et al, [11] take a statistical approach to find the factors affecting technology integration in K-12 education. Teachers' readiness and school level factors (availability of computers, technical support and overall support) positively influenced technology integration, while teachers' demographic characters (age and years of teaching) negatively influenced technology integration in Schools.

On a more specific level, [12] discuss the factors that influence the use of Tablets in a K-12 classroom setting. The biggest one being availability of a smooth running technical infrastructure and support system. Similarly [13] discusses integrating laptops in K-12 classrooms in which the results suggest that teacher readiness and teacher beliefs strongly predict laptop integration, and that overall support for school technology and professional development have strong effects on teacher beliefs and readiness, respectively. According to a report [1] on future trends in educational technology, gesture based technologies will become more commonplace in classrooms. The purpose of the study is to provide design guidelines for the integration of such gesture based systems in classrooms.

2.2. Commodity motion capture systems

This section presents some of the motion capture systems that are available in the market. Optical motion capture systems can be broadly classified into marker-based and markerless motion capture systems. There are many other types (inertial, magnetic and mechanical) of motion capture systems available in the market, but we are only interested in these two because they are widely used. Research papers comparing these two technologies tend to focus more on the tracking accuracy and tracking algorithms used [14, 15]. But this thesis is geared towards designing the system for a classroom environment.

Marker based motion capture system

Marker based motion capture systems are popular among animation and game studios for their accurate tracking. It usually consists of several high resolution optical infrared cameras mounted on a wall or a tripod. Figure 1 shows an Optitrack marker based motion capture system which has 4 tripods and 8 cameras mounted. These cameras detect the 3D positions of markers which are present on the actor. The markers can be active or passive. Active markers are usually made of LEDs that can emit infrared light and passive markers are made of retro reflective materials that reflect infrared light falling on them.

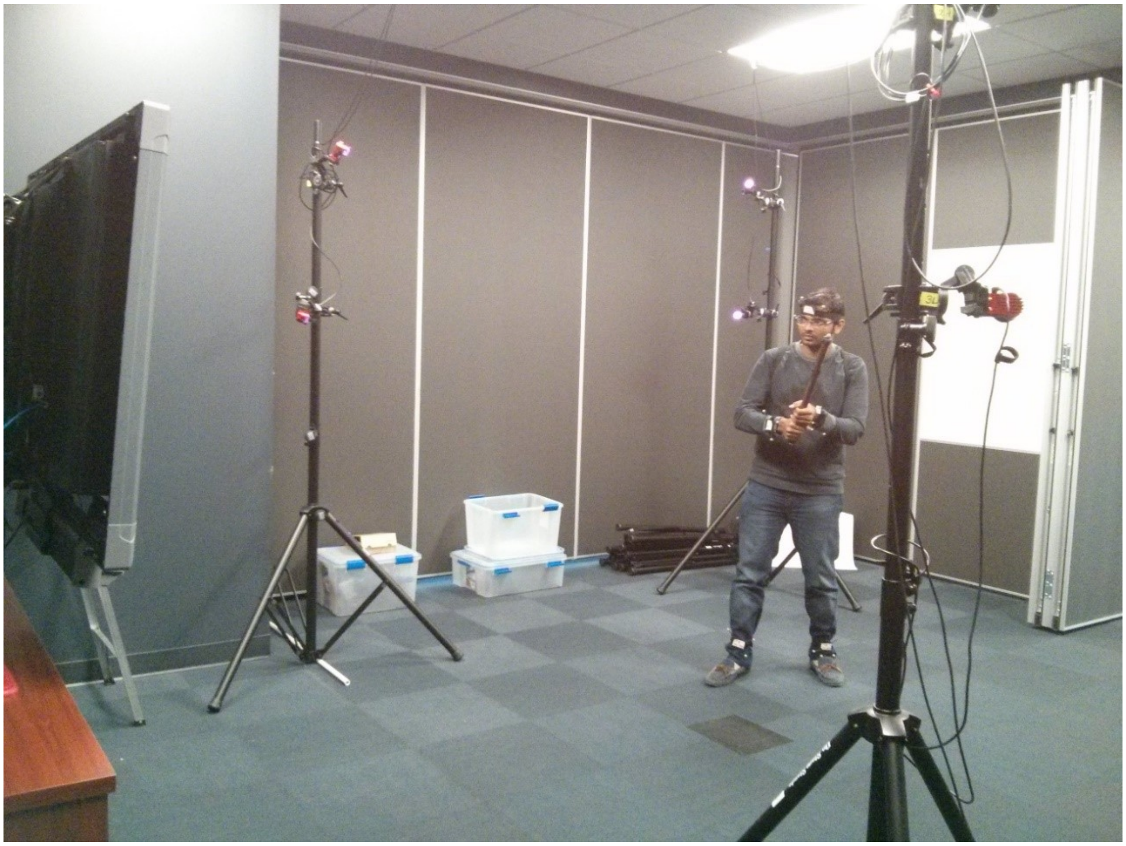

Figure 1 : Optitrack motion capture setup with Flex 13 cameras

The multi view data obtained from the cameras need to be calibrated to obtain positions of the cameras. Calibration is a three step process which includes dynamic, static and skeletal calibration. Dynamic calibration involves waving the 3 marker calibration wand in the capture volume. The exact length of the calibration wand is known to the calibration engine and that information is used to analyze thousands of samples collected during dynamic calibration. This calibration is needed to calculate intrinsic and extrinsic parameters of the cameras. Static calibration is a step where an L-shaped calibration square is placed on the ground to determine the global coordinate system. Skeletal calibration involves placing a set of markers on the actor's motion capture suit and registering that with the software in a particular order. Whenever the complete skeletal tracking is not required, it is possible to create a set of rigid bodies which consists of at least 3 markers.

After the calibration is complete, the data can be streamed to client applications over the network. Calibration processes for these cameras are usually non trivial and require dedicated technical personnel and expensive software. This is feasible in a studio or a research lab, but it is difficult to imagine running such a system in a school setting.

Markerless motion capture system

Markerless motion capture systems use emerging technologies for tracking. They do not involve the hassle of putting multiple markers. One of the most popular markerless motion capture system is the Microsoft Kinect [16]. The applications of the Kinect range from animating virtual characters [17] to realistic full body 3d reconstruction [18]. Global Kinect sales had passed 24 Million [19] in 2013 and many Schools have integrated Kinect for experiments in classrooms. Kinect also comes with an SDK that lets you stream the skeletal joint coordinates to any application through USB port.

Figure 2 : Microsoft Kinect for Windows v2 sensor

In 2014, Microsoft released the second generation model of the Kinect (shown in figure 2) with significant improvements compared to its previous model [20]. The latest sensor can now track up to six skeletons (previously limited to two) and 25 joints per person. With improved color camera resolution, depth fidelity and wider field of view, the capture volume of the sensor is increased. Kinect is not only convenient to use but also provides the quality of tracking performance which is needed for our purposes.

2.3. Comparison of marker-based and markerless motion capture systems

There have been many papers that compare the quality of the motion capture data between marker-based and markerless motion capture systems. Puthenveetil, S.C., et al [14] discuss the operational principles and the tracking accuracy of the two systems. They are interested in the accuracy of the body joint angles measured by the two systems when a person is performing a fastening operation on a physical mockup of an aircraft fuselage. They showed that even though the quality of accuracy of the Kinect is lower than that of the marker based system, it can still be utilized for capturing simple human movements in industrial engineering applications. Chang, C.Y., et al [15] compared the two systems as a tool for virtual reality rehabilitation. They showed that the Kinect has the potential to be a rehabilitation intervention tool by conducting a user study with participants with spinal cord injury.

Even though skeletal tracking accuracy of the Kinect can be compared to the accuracy of Optical marker based solutions like Optitrack or Vicon, its object tracking is not very reliable. Object tracking solution [21] for the Kinect usually involves a deterministic search of the Kinect RGB image from its camera, with a reference image. But these solutions expect the background to be uncluttered and not contain objects with similar color as that of the reference image. Dutta T., [22] provided some details on the marker design guidelines for doing object tracking with the Kinect. He acknowledged that the object detection was very difficult when the object was highly reflective or absorbed light. The probability of object detection decreased as the object moved away from the sensor. Ren, H., et al [23] developed a proof of concept attempting to track four retroreflective markers just by using the depth images from two Kinect devices in sub millimeter accuracy for surgical applications. Han J. et al, [24] discuss different algorithms for object recognition and object tracking: it indicates that the data processing load of using RGB and depth image is so high that it is impractical for real-time applications. In this research we leverage the research conducted on the accuracy of the motion capture systems to design a pragmatic solution for a classroom setting. The next section discusses some applications of motion capture in education.

2.4. Motion capture systems in education

This section presents some of the examples of using motion based technologies to accomplish educational objectives. The educational foundation of using motion based technologies is based on the idea that there are three types of learners, auditory, visual and kinesthetic. Kinesthetic learners learn better when they touch or are physically involved in what they are studying. They constitute 15% of the population[25] and it is important to cater to these students while developing teaching methods. There is rich research on the significance of gestures in education. Alibali, M.W. et al [26] presented evidence drawn from teachers' and learners' gestures to suggest that the mathematical knowledge is embodied. Cook, S.W., et al [27] have shown that gesturing during speech can lead to better recall.

Motion capture based technologies can be used to develop embodied learning experiences. Embodied learning combines human computer interaction and cognitive science to create interactive educational experience. Research has shown compelling evidence that nearly all of our experiences are grounded in the body. Johnson-Glenberg, M.C., et al [4] discussed an embodied mixed reality learning environment called EMRELE that showed significant learning benefits when compared to regular classroom instruction. They attribute the learning gains to the level of embodiment in the lessons. They also propose a taxonomy of embodied learning in educational spaces in which the highest level of embodied learning (fourth degree) involves three components: sensorimotor activation, gestural congruency with content and perception of immersion. Sensorimotor skills involve the process of receiving sensory messages (sensory input) and producing a response (motor output). We receive sensory information from our bodies and the environment through our sensory systems (vision, hearing, smell, taste, touch, vestibular, and proprioception). Even though this is an active area of research, it is still not very clear how exactly kinesthetic experiences improve learning [5].

Marker based technology was used in [3] to improve the quality of physical education by detecting body gestures. In [28] marker based motion capture system was used to create a virtual reality dance training system, in which the participants showed that they can improve their dancing skills using the proposed system. Vrellis, I. et al [8] studied the attitude of primary school children towards the Kinect and Mouse. The Kinect was preferred over the Mouse even though it was less user friendly. [5] lists some of the applications created with the Kinect that has the potential to impact education. But the author acknowledges the technical and pedagogical constraints with respect to integrating it in a classroom setting, indicating that more research is necessary to address some of these problems. More solid empirical evidence and inter-disciplinary work is needed to do this. Hence this thesis aims to inform developers and researchers in psychology, education and virtual reality to design a practical motion capture system.

2.5. Digital Micro Enactment (DiME) marker based system

This section presents a system called DiME that was proposed by Chu et al [2] and used as an example to test the motion capture design criteria. This system was chosen because the primary goal of the system which is storytelling is extensively used in elementary school curriculum. Also DiME marker based system was successfully tested in an after school program by Chu et al [2] and showed that the motion based technology can improve children's sense of self efficacy. DiME system can create animated stories using marker based motion capture system.

Animated stories are one of the powerful ways of storytelling that can captivate children of any age or culture. There are many instructional design experiments conducted that have shown that Digital Storytelling can increase learning motivation, critical thinking and nurture creativity [29]. But animation creation is not an easy task even for adults. It is an ongoing effort to create animation creation interface for kids. There are some approaches that allow manipulating real world objects to drive animations and others where a traditional keyboard and mouse interface are used to drive animation.

DiME marker based system captures body gestures of the child and movement of a physical object held by the child and mirrors that onto an animated character and a virtual prop as shown in figure 3. Children work in a group of two to come up with embodied, story fragments called micro enactments [30], which are later put together to create an animated story. DiME is based on Performative authoring concept [31] which uses the power of pretend play to create animation at real time and has been successfully tested in Schools. The target users for this system are children of age 8-10. In psychology, it is known that children undergo a creativity slump during this period. Piaget calls this the concrete operational phase when their thinking becomes more organized and rational. The goal of this system is to involve children in the 'process of creativity' through micro enactments. Children can use a variety of physical objects to enact their stories, which are converted to virtual objects in the story scene. Optitrack motion capture was used for this study and hence this system will be referred to as DiME marker based system.

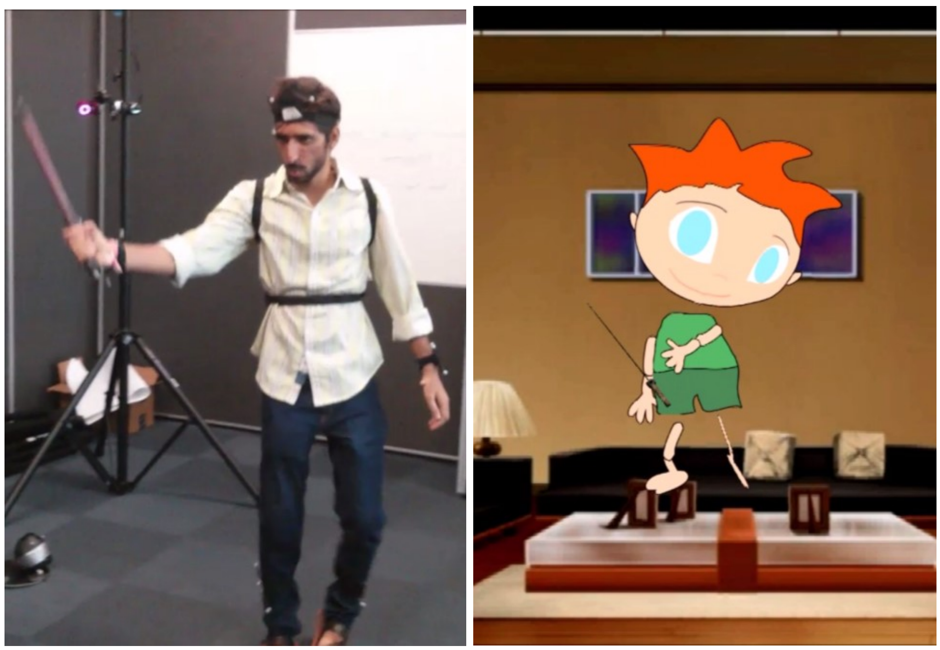

Figure 3: Body gestures made by the actor holding a physical object (Left), animated character with a virtual prop (right)

Optitrack motion capture system with 4 Tripods, 8 infrared Flex 13 cameras was used for motion tracking. The system costed about $17,800 at the time of purchase. This system was successfully tested in a local after school program in a month long user study. Since this is a marker based system, children wear 18 markers (6 for hands, 6 for legs, 3 each for head and torso) which are tied using Velcro bands. Usually the calibration process takes about 10 minutes to complete using a standalone software. The markers are created using a software called Motive which is provided by the same company that makes Optitrack system: Natural Point. We used this system as a probe to get opinions from the teachers about possibility of classroom integration of DiME.

3. METHODOLOGY

There are 3 stakeholders in a typical school ecology: the children, the teachers and the administrators. Inan et al [11] have shown that the teachers play a significant role in the technology integration. Since we are interested in the classroom integration of the technology and not testing a specific system, we decided to interview elementary school teachers and not the children. Teachers also provided some input about the perceptions of the children and the school administrators during these interviews.

A three step research approach was followed in which DiME marker based system was used as an exemplar system to answer the research question of designing the motion capture system for the classroom. In the first stage, a formative study was conducted with school teachers to determine the requirements and constraints for DiME marker based system. Design criteria were formulated based on the findings of this study. In the second stage, the design criteria was used to design and implement the DiME markerless system. In the third stage, an evaluative study with elementary school teachers was conducted to qualitatively and quantitatively compare DiME marker-based and DiME markerless system. Nine design guidelines emerged as the themes of this user study, which are discussed in a later section.

3.1. Formative study

The purpose of the formative study was to present the DiME marker based system to the teachers and get an idea about the feasibility of motion capture technology by learning more about current technologies used in the classroom. A user study was conducted with a focus group of 4 teachers to understand how the DiME marker based system can be integrated in classrooms. The semi structured focus group was conducted on the university campus in a lab setting which lasted for 2.5 hours. The teachers were recruited through emails to College Station Independent School District and Bryan Independent School District. The teachers were asked to sign and return a consent form to audio and video record the study, through email before participating in the study. The discussion topics covered the following areas.

1. Introduction and questions on teaching experience
2. Current methods of teaching storytelling in elementary schools
3. Technologies (Hardware/Software) used during teaching
4. Presentation of DiME marker-based system
5. Feedback about the possibility of classroom integration
6. Feedback about the usability of the system
7. Feedback about DiME markerless (early prototype)

3.2. Findings of the study

About 3 hours of video data was collected and transcribed using Inqscribe software to perform qualitative coding. Open coding process was performed on the transcript and then it was grouped into categories. Selective coding was then performed to uncover main themes of the discussion from these categories. The teachers were given code names (T1 to T4) and their experience with elementary school children is as follows. Also T2, T3 and T4 taught at the same school. Teachers had a wide range of experience from 5 years to 30 years and had handled subjects like math and English language arts (ELA).

Table 1: Teacher demographics for the formative study

Teacher code	Years of experience	Subjects taught
T1	30	Math
T2	13	ELA
T3	7	ELA, Reading specialist
T4	5	ELA

The main themes that emerged out of the discussions are summarized below.

Current technologies (software/hardware) used in classroom

T2-T4 mentioned some of the apps that they extensively used in their classes. One of them was *Spelling city* that they used to work on spelling and vocabulary using interactive games on their iPads. Teachers considered it very useful because it also indicated the problem areas, frequently misspelled words of each child, which is very difficult to track for them using paper based tests. They also said that they saw a significant improvement in their grades after using that app since they can ask them to practice the most misspelled words. T2 said "*they (students) are so into it. For me I can track their progress. If they have their most misspelled words I can give that to them and they keep practicing those*."

Teachers mentioned another app that shares what they type or write on the smart boards. This way the teachers can ask a question and the kids can draw/type their answers and see them on the big screen. Teachers mentioned that they look for apps that are available for free. They usually need to write a grant proposal in order to purchase an app. One of the teachers also mentioned about calibration that is needed for Smart boards which takes about 20 to 30 seconds to complete. Teachers are used to performing this simple calibration process.

Feedback about integrating storytelling system in curriculum

Teachers gave a positive feedback about the concept of performative authoring and the idea of using motion based technologies in their classroom. T4 said *"I am already getting the idea. I agree with the writing that you act out and write stories but it's the same thing with the reading like if we were to read a class novel together if students pick up a scene from that novel and re-enact to show their comprehension of the book."* Interestingly, all the 4 teachers mentioned that they would use the system differently. While the language arts teachers said that she would make the students act out parts of a novel that they read, the Math teacher said that she could use multiple props to convey addition and multiplication. Since all the teachers thought that the DiME system could be used to match their learning objectives, it also validates the decision to choose DiME system as an example to test the use of motion capture systems.

Administrative constraints of integrating DiME marker based system

These are some of the concerns raised by the teachers that are beyond the control of the teachers and are best addressed by the school administrative authorities. Factors like classroom space and cost of the system are usually addressed by the decision makers in the respective school.

Space constraints for the Tripods

Teachers thought that the tripods and cables occupied a lot of classroom space. A large capture volume means it needs a dedicated floor space apart from very high costs associated with the cameras. T1 summarized this as *"I don't have that much time to spend. Take the kids up, go down to the lab to do you know"* T2 said *"I would mount the cameras in the classrooms on the walls, they are going to kick the tripods and they can't take up too much space"*

Cost

Many Schools today have access to Interactive White Boards (price ranges from about US $800 to US $2,500) [5] but the cost of the marker-based motion capture system is more than 10,000$ which is beyond the budget of many schools. The teachers thought that having a system which costs more than 10,000$ in every classroom was very unrealistic. In the two schools that the teachers came from, they relied on writing a grant proposal to acquire new technology into the classrooms. *"We spent couple of hundred dollars for spelling city and 500$ for the robotics kit" said T3.*

Portability

Teachers mentioned that usually all the technology starts small in the school with just one teacher using it. After the technology proves itself, it becomes more prevalent. So it is common to share the resources among the teachers. So they wanted the system to be usable by different teachers for their classes easily. According to the teachers that were interviewed, the Tablet computers used in schools were swapped and shared with multiple students. Teachers said that the ability to check out the system for a certain class and release the resource for another teacher to use would be beneficial. “See portability is going to be huge” said T2.

Technical constraints of integrating DiME marker based system

Apart from the administrative constraints, the teachers indicated that the DiME marker based system needed a lot of work from the teachers to setup. These include understanding the calibration software, setting up the system and placing the markers on the actors. Teachers who use a lot of technology in their classroom are usually more open to learning these things than others. But the ideal system should abstract the technical details of the system from the teachers.

Technical expertise

Marker based systems like Optitrack usually relies on a proprietary calibration engine that comes with another software package. DiME marker based system needs a software called ‘Motive’ to be running in the background to track the markers. Teachers thought that it would be a steep learning curve to learn this software and understand rigid body calibration. Teachers indicated that they cannot spend a lot of time learning new technology. Setting up an Optitrack system requires familiarity with 3D navigation. *"It takes half a day to setup! it's an event, then it has a limited application" said T1.*

Calibration process

The calibration process takes at least 5 minutes to complete and is difficult to be performed by non-experts. If the tripods slightly move or the markers fall out, then it will require recalibration of the system. This was one of the comments by the Teachers. *"… it (DiME marker based system) has to be user friendly. So that the people are not going "I dont have time to learn that"*

Marker placements

Existing setup consisted of placing 18 markers on the actor, which would take up a lot of class time to take the markers on and off. Teachers also suggested that the Markers looked delicate and fourth graders could break them easily.

Feedback about early prototype of DiME markerless

An initial prototype application was developed using Microsoft Kinect markerless system with only the skeletal tracking functionality to get feedback about the system. The teachers had a positive response towards the usability of this system.

T3 said *"Yes this one does because I think Schools would buy one of this and kind of test it out first and let's say third grade is going to test it I could use it one day and take it to T4 and she could hook it up and use it". "This is something, the portability of this is something that could be used."* said T2. These themes in the formative user study were used to formulate a list of design criteria that are required for a motion capture setup to work in a classroom.

3.3. Design criteria

Themes found in the section above can be summarized as in the table 2. The goal of the rest of the thesis is to design a system that addresses these issues with the DiME marker based system.

Table 2: Design criteria for motion capture systems

Constraint	Design Criteria	DiME marker based
Administrative	Space	12ft x 12ft
Administrative	Cost	>10,000$
Administrative	Technical complexity	Learning Motive software
Technical	Marker placement	18 markers
Technical	Calibration time	5-10 minutes
Technical	Portability	No

There is one more constraint that was added to ensure the motion capture systems are comparable. The definition of motion capture is that it can track the movements of objects or people. Most marker based systems have the ability to perform both object tracking and skeletal tracking. Since we are interested in coming up with guidelines that are generalizable, it is important that both the systems are comparable. So the new system need to have the ability to track both objects and actors.

4. IMPLEMENTATION

4.1. DiME markerless system

Based on the design criteria introduced in the previous section, DiME markerless motion capture system was implemented. The features of this system is the same as the DiME marker based system except, it uses a different set of hardware and offers a different calibration method. Figure 4 shows the schematic diagram that points out the difference between the systems. It makes use of commodity gaming hardware available in the market and only requires about 2 square feet area for the hardware. DiME markerless system uses a Microsoft Kinect and Sony PlayStation Move sensors for tracking and the setup is shown in figure 5.

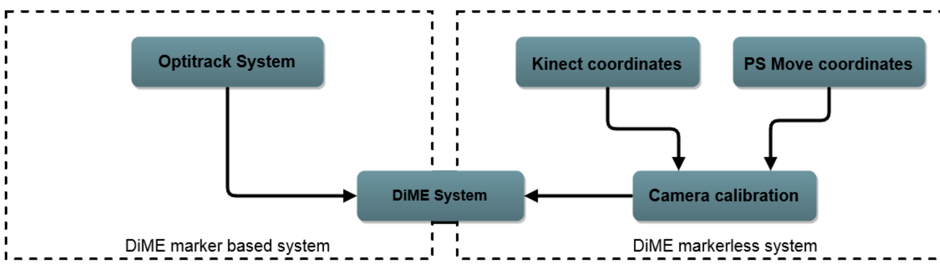

Figure 4: Motion capture input for DiME marker-based and DiME markerless

Figure 5: DiME markerless system hardware

4.2. Choice of hardware

Object tracking and skeletal tracking are crucial for the DiME system. The Kinect version 2 has made the system very robust for skeletal tracking which is also provided with Kinect for Windows SDK 2.0 API. But the Kinect camera has limitations in detecting objects. Using retroreflective markers with Kinect depth sensing is a possibility but it is known to be problematic for object recognition[22] and not supported by the API. In theory it is possible to attach accelerometers to the actors and use that information to track animated objects as suggested by [32], but these are susceptible to drift over time and are quite noisy. This begs for more robust object detection techniques using accelerometers or inertial measurement units (IMU) like PlayStation Move or Wii controllers.

There are some options available for 6 DOF object tracking like the PlayStation Move, Oculus Rift, Razer Hydra and Nintendo Wii motion Plus. Oculus rift which has gained a lot of attention recently has an age limit of thirteen [33] and does not help the age group that we are targeting. Other devices like Razer Hydra uses a magnetometer to perform object tracking in 3D space but it has a very limited capture volume compared to the Kinect and PlayStation Move. Nintendo Wii motion plus is less accurate when compared to PlayStation Move, since the PlayStation Move also has a magnetometer.

APIs for Kinect and PlayStation Move uses two opposing strategies. While the Kinect offers an open source SDK for using the sensor, the PlayStation Move can only be used by using it along with a PlayStation 3 console hooked to the same network. This could also be the reason why the former is so popular with the developers [34]. In order for the PlayStation Move to be tracked by a client application, a server application called Move.me needs to be running on a PlayStation 3 console on the local network. This adds an additional cost of about $250 for PS3 console apart from the controller which costs about $50. Also Move.me server application is available on the PlayStation network for $100. An attempt was also made to track object without using the console as specified in [35]. But this solution is known to be problematic in Windows, which is the required platform for the Kinect.

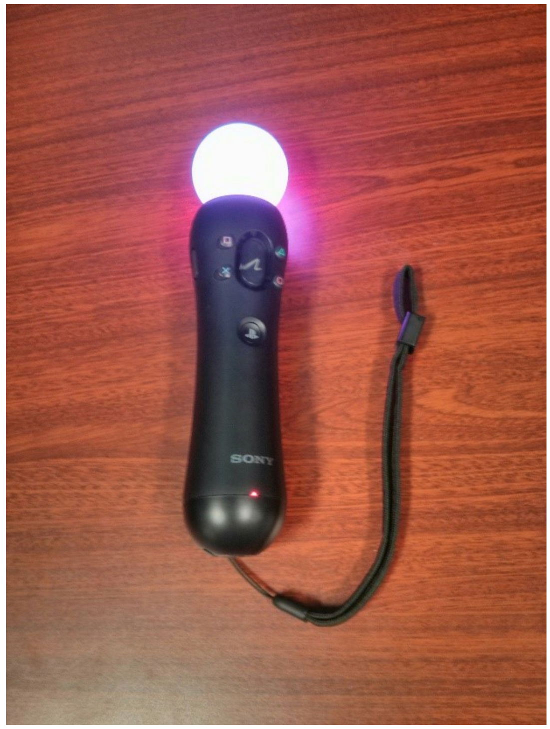

Figure 6: Sony PlayStation Move controller

PlayStation Move (PS Move) is a motion based controller for the PlayStation 3 (PS3) console by Sony. It provides 6 degrees of freedom tracking (position and rotation in 3D) using the glowing blob on top of the controller. Inspired by the success of the Nintendo's Wii console, PlayStation Move uses the PS Eye camera to track the wand's position, and inertial sensors to detect its motion. This is a hybrid motion capture system where Optical motion tracking is used for detecting the position of the wand and the magnetic and inertial motion capture system is used for rotational tracking. The glowing blob consists of 3 LEDs of red, blue and green that can be set by an external application.

4.3. Architecture overview

Unity game engine was considered as a feasible option for developing DiME markerless system. Unity is compatible with a variety of motion sensing hardware like the Kinect, Oculus rift and Razer Hydra. It also follows "author once, deploy everywhere" principle, so that the same application developed can be deployed on Windows, Mac, Linux, XBox or even Mobile platforms. The game engine can also be used in both the Mac and PC operating systems and it uses the JavaScript and Just-In Time compilation within the C++ mono library. It also uses the Nvidia PhysX physics engine, OpenGL and DirectX for 3D rendering and OpenAL for audio.

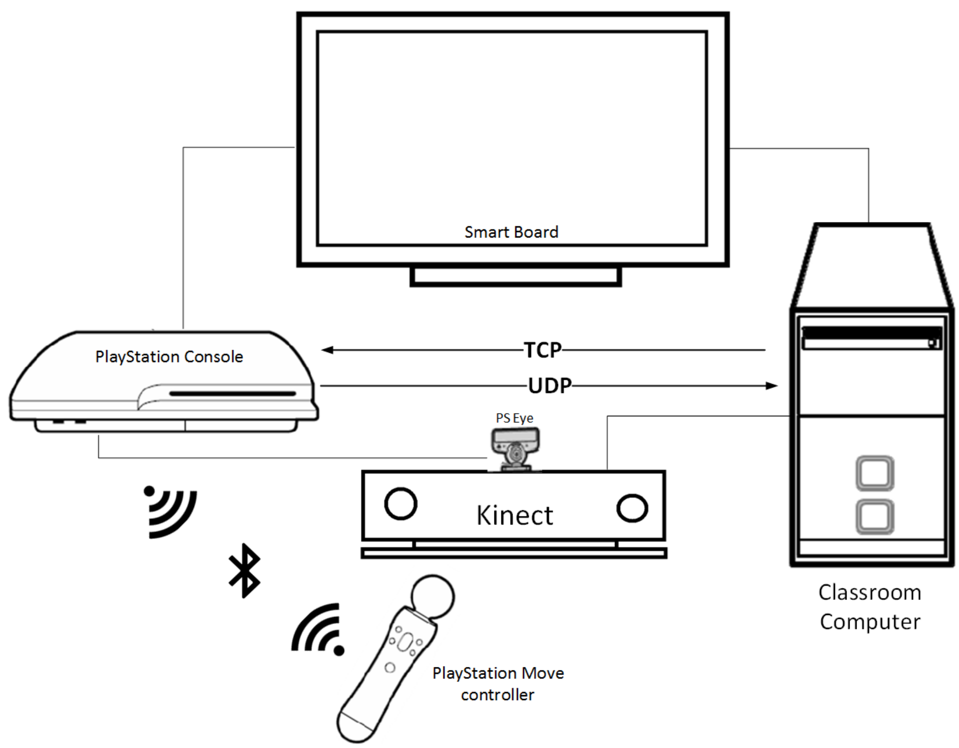

Figure 7: DiME markerless system hardware setup

Figure 7 shows the hardware setup of DiME markerless system. The PlayStation controller communicates with the PlayStation console through Bluetooth. The controller can be charged through USB. The Move.Me server application on the PlayStation console accepts connections over TCP on port 7899. It supports up to four simultaneous connections from the clients. Once a client connects through TCP, it sends an initialization command to the server along with a UDP port for the server to use when sending data to the client. After this stage, the server sends the current state of all of the motion controllers through UDP to the client application. Kinect is supported through Kinect for Windows SDK 2.0 which works on Windows 8 or above. Also Kinect needs to be connected to a USB 3.0 port on the development PC.

4.4. Calibration technique

In the formative user study one of the teachers had mentioned that the calibration for Smart boards are trivial and can even be done by the students. Smart board calibration usually takes about 15 seconds. The calibration technique presented here takes 2 samples per second for 20 seconds from the Kinect and PlayStation sensors. The users are encouraged to make calm motion with the controllers on their right hand. It is assumed that the position of the center of the controller returned by the PlayStation wand and the position of the right wrist of the Kinect are the same. This will introduce minor systemic error but this assumption is crucial for the simple calibration. These devices are assumed to be calibrated correctly and the rigid body transformation is calculated which aligns PlayStation coordinates to the Kinect coordinate system. The technique used in this system are generalizable to any other motion capture based application which is to be deployed in a classroom setting.

There are many widely used calibration and registration techniques in computer vision. We are specifically dealing with registering 3 degrees of freedom(DoF) input dataset which we get from Microsoft Kinect and Playstation Move. Data is composed of x,y,z values from two different coordinate systems. We take n corresponding samples from both the systems in T seconds. Figure 8 shows the samples taken during the calibration from DiME markerless system. The Pink spheres indicate the samples that are taken already.

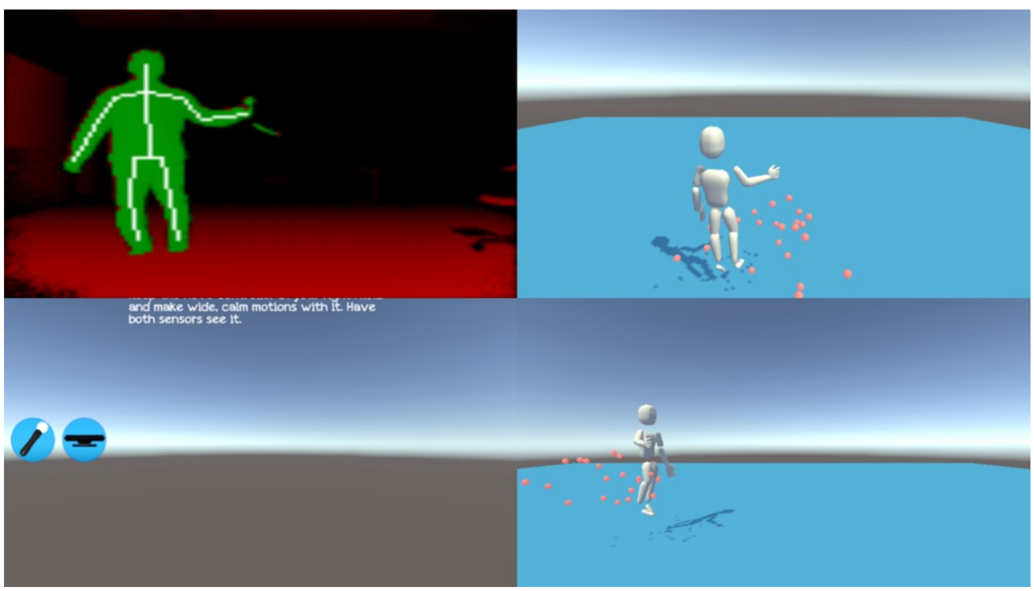

Figure 8: Samples taken with the PS Move controller on the right hand

Samples from the Kinect and the PlayStation are stored in two matrices A and B.

$$A = \{a_1, a_2, a_3, \dots, a_n\}$$

$$B = \{b_1, b_2, b_3, \ldots, b_n\}$$

Such that

$$\boldsymbol{b_i} = \boldsymbol{R}\boldsymbol{a_i} + \boldsymbol{t}$$

where R is a 3x3 rotation matrix and t is the translation vector. This is a system of overdetermined equations where there are more solutions than unknowns. We know that the samples obtained from the sensors are prone to errors. This becomes an optimization problem where we need to minimize the sum of squares of those errors. Mathematically we can express this as an optimization problem

$$\min_{R,t} \sum_{i=1}^{n} ||Ra_i + t - b_i||$$

One of the standard ways to solve this optimization problem is using Moore–Penrose pseudoinverse matrix[36]. Let X be the transformation matrix that we need to find. This problem is of type

$\boldsymbol{AX} = \boldsymbol{B}$ which can be solved by using pseudoinverse matrix as

$$\boldsymbol{X} = (\boldsymbol{A^t}.\boldsymbol{A})^{\boldsymbol{\iota}}\boldsymbol{A^t}\boldsymbol{B}$$

X is a 4x3 transformation matrix and we can extract the rotation R matrix and translation t. These values are stored in the application and works well until the camera positions don’t change. There are many other ways to solve this optimization problem which are compared in this paper[37]. In this system, the solution can be found with 40 samples that taken over 20 seconds.

4.5. Comparison between the DiME systems

Table 3 shows the quantitative comparison between DiME marker based system and DiME markerless system. While the design criteria has yielded a better system in terms of all these aspects, this system must be validated with a evaluative user study which will be addressed in the next section.

Table 3: Comparison of the design criteria for motion capture systems

Constraint	Design Criteria	DiME marker based	DiME markerless
Administrative	Space	12ft x 12ft	2ft x 2ft
	Cost	>10,000$	<500$
	Portability	No	Yes
Technical	Marker placement	18 markers	Markerless
	Calibration time	5-10 minutes	20 seconds
	Technical complexity	Learning to work with a 3D software	Additional software not necessary

5. EVALUATION

5.1. Research questions

An evaluative study was conducted to measure the effectiveness of the design criteria introduced in the previous section. This study was designed to answer the following research questions.

1. How effective are the design criteria for classroom motion capture system?
2. What are teachers' assessment of the usability of DiME marker based system and DiME markerless system?
3. How can the affordances of motion capture system help to accomplish learning objectives in classroom teaching?

Answering these questions would result in the design guidelines for the researchers and developers in educational technology to design motion capture systems for classrooms. It would provide some quantitative data about usability of marker based and markerless system. It would also inform about the teachers' attitude towards integrating motion based technology in classrooms.

5.2. Study design

A user study was conducted by interviewing 6 teachers to understand how DiME markerless system compares to DiME marker-based system. System Usability Scale (SUS) was used to quantitatively measure the usability of the two systems. Additional questions were asked about the system to get more insights about feasibility of this ecology of devices in the classroom. Finally teachers were asked to write lesson plans that used motion capture system to accomplish their learning objectives.

The interview was conducted on the university campus and lasted for about 1.5 hours each. The qualitative and quantitative data obtained are discussed in this section. The teachers were recruited through emails to College Station Independent School District and Bryan Independent School District. The teachers were asked to sign and return a consent form to audio and video record the study, through email before participating in the study. The interview protocol is presented below and the detailed protocol can be found in Appendix 2.

1. Introduction and questions on teaching experience
2. Presentation slides about the concept of DiME
3. Presentation slides about the DiME marker based system
4. Qualitative questions
    - Would it be practical to use this System in School as is? What changes do you think are necessary?
    - Could you imagine teaching your class by using this System?
    - What did you find confusing in the System?
    - What did you dislike about the System?
    - What did you like about the System?
5. SUS Questionnaire
6. Lesson plan

The steps 3 to 5 was repeated for DiME markerless system. To prevent the effect of the order of presentation, the participants were divided into two groups and the order was reversed for the second group.

About 9 hours of video data was collected and transcribed using Inqscribe software to perform qualitative coding. Open coding process was performed on the transcript and then it was grouped into categories. Selective coding was then performed to uncover the main themes of the discussion. The teachers were given code names (T1 to T6) and their experience with elementary school children is as follows. It is also worth noting that T5 and T6 were also present for the formative user study.

Table 4: Teacher demographics for the evaluative user study

Teacher code	Years of experience	Subjects taught
T1	3	ELA
T2	14	ELA
T3	30	ELA
T4	4	Math, Science, Social studies
T5	30	Math
T6	13	ELA

5.3. Quantitative feedback

System Usability Scale (SUS) [38] was used to test the usability of DiME markerless and DiME marker based systems. SUS consists of 10 statements that are scored on a 5-point likert scale based on the strength of agreement. The questionnaire used in the study can be found in Appendix. SUS is generally seen as providing a high level, subjective view of usability and is often used to compare two different systems. Bangor et al[39] described 2324 survey results from 206 usability tests over a ten year period and showed that the SUS was very reliable over a wide variety of interface types.

The final score of SUS ranges from 0-100, where higher score indicates better usability. SUS scores are not to be confused with Percentage. The average SUS score from 500 studies is 68 and anything below 68 would be considered below average. Figure 10 shows the overall trend of the final SUS score given by each participant which clearly shows a positive valence towards DiME markerless system. The graph clearly indicates that all the teachers thought DiME markerless system was more usable. Also the average SUS score for DiME marker based was 41.67 and for the DiME markerless was 75.42.

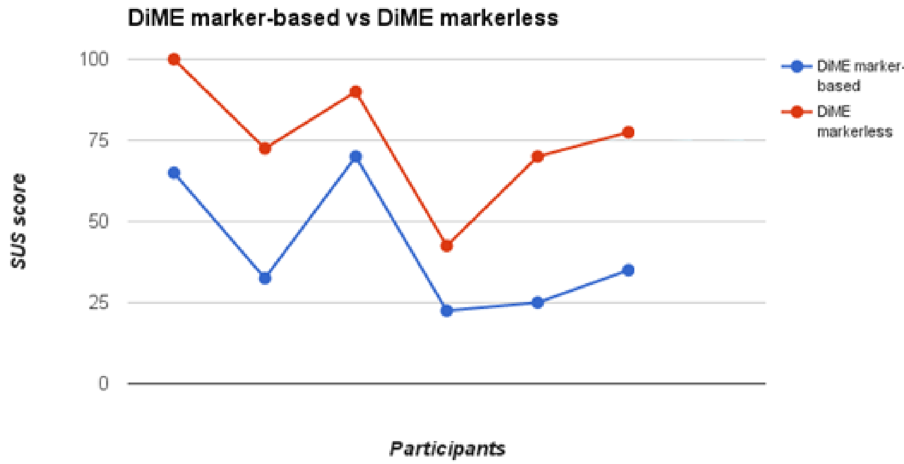

Figure 9: SUS scores for DiME marker based and DiME markerless for 6 participants

Since the sample size of our data was only 6 and the data did not follow normal distribution, the Wilcoxon-Mann-Whitney test was used on the 10 statement scores and on the average test score. The Wilcoxon-Mann-Whitney test is a non parametric statistical hypothesis test used when comparing two related samples to predict whether their population mean ranks differ. Table 5 shows the variables that showed statistically significant ($p < 0.05$) differences between marker based and markerless systems.

Table 5: Statistically significant statements from SUS

Metric	Significance (Z, p <0.05)
I think that I would like to use this system frequently	(-2.041, .041)
I thought the system was easy to use	(-2.032, .042)
I think that I would need the support of a technical person to be able to use the system	(-2.264, .024)
I would imagine that most people would learn to use this system very quickly	(-2.271, .023)
I found the system very cumbersome to use	(-2.264, .024)

5.4. Lesson plan

All the teachers were asked to write out lesson plans that would use DiME system. The most common use of DiME system was by ELA teachers to write out personal narratives which is a big part of STAAR (State of Texas Assessments of Academic Readiness) test. While T6 said that she could use this system to teach ELA every day, T3 and T1 said that they might use it every other day. T5 said that she will not be able to use the storytelling system to teach Math. T4 said that he could use the system in social studies to re-enact scenes that occurred in Texas history. This suggests that integrating DiME markerless system in classrooms can provide a lot of learning benefits and augment classroom teaching.

5.5. Qualitative feedback

The main themes of the data encoding are presented in this section. Some of the themes that were discussed in the formative study were repeated here. Based on the number of teachers who mentioned the design criteria in their comments, the design criteria were validated to form the design guidelines discussed in the next section. The main themes of the study are discussed below.

DiME marker based system

Overall the teachers had many issues with the DiME marker based system which is also reflected in the quantitative findings. T2 and T3 thought that the markers were too fragile and that the students would break them easily. T5 mentioned that the additional software Motive that needs to run in the background, needs to be shown to a technology specialist on campus, to understand if it is possible to run it on the classroom computers. T4 thought the system was not safe for fourth graders who are running around the class because of the cables and the tripods involved. T2 summarized all the feedback as *"... Having all the cameras and all their components, cost and wear and tear with the kids. There should be little or no assistance from me. But in this system I think I would constantly need to help them"*

DiME markerless system

Teachers had a positive outlook towards DiME markerless system. 4 of the teachers said that they could see themselves using the DiME markerless system in their classrooms. The graph shown in figure 9 shows the number of teachers who thought it was practical to use the system in their classroom. It is interesting to note that the teachers who said yes were all ELA teachers. (T1, T2, T3, T6) were all ELA teachers.

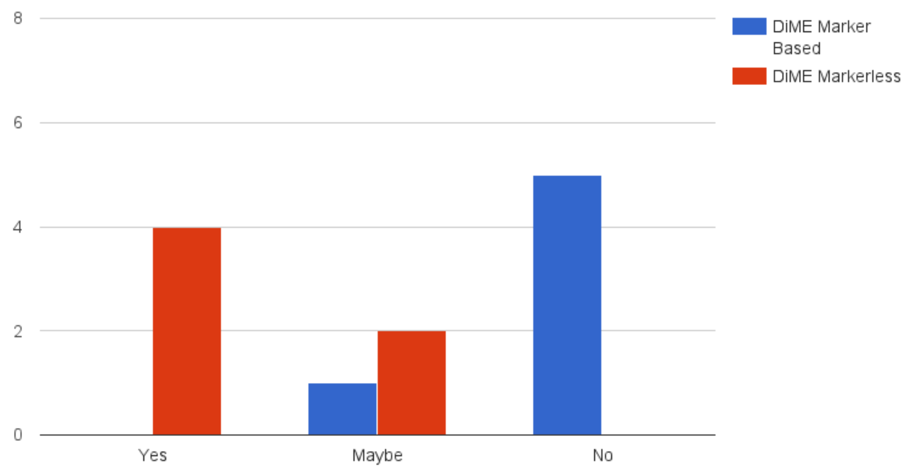

Figure 10: Teachers' response to whether they thought it was practical to use DiME system

Themes related to the design of the system

Many themes related to the design of motion capture systems got repeated and were classified. 3 more themes were noticed in this study which belongs to the pedagogical constraints that the teachers face. The frequency of the repetitions are listed in table 6.

Table 6: Design guidelines along with the frequency of repetitions

Constraint	Design Guideline	Number of teachers who mentioned this theme in the interview (Total = 6)
Administrative	Space	5
	Cost	5
	Portability	3
Technical	Marker placement	4
	Calibration time	2
	Technical complexity	6
Pedagogical	Student engagement	3
	Curriculum match	6
	Familiarity	4

Space

Two of the teachers said that the classroom space in their schools are still not enough to incorporate DiME markerless. But three others said that this can be setup in their classroom. This is how T6 said that she would setup in her classroom – *"I would want to put the equipment there (PlayStation console and Kinect sensor) on a wall mount so that the kids can't touch it. I would put it right above my smartboard, so that's the area to enact, and if I use it like a station or something where the kids could go and* *enact". "…. you can setup in half of a classroom or corner of a library."T2 on DiME markerless.*

Cost

The combined cost of the hardware needed for DiME markerless system is about 500$. Teachers thought this was well under the budget of schools nowadays. T1 mentioned that the teachers can write a grant and get this system in their classroom – *"We could write a grant and get this one on Monday."*

Portability

DiME markerless system can easily be moved in and out of a classroom, since it does not occupy a lot of space. This allows teachers to share the resources among each other. *"a portable thing that you can setup in half of a classroom or corner of a library" – T5 on DiME markerless*

Calibration process

The process of calibration was improved significantly in DiME markerless system which the teachers thought was convenient. Two teachers talked about the process in their comments and they both thought that the kids would manage to perform calibration process easily. T6 said "*I mean calibration wasn't confusing, I mean how often do we have to calibrate our smartboards, I mean it's just something you do. That to me was very simple. Just a matter of moving it, kids can totally do that."*

Markers

DiME markerless system uses markerless approach for skeletal tracking which obviates the hassle of wearing markers on the actors. The actors can easily be swapped during enactment. "*you don't even need to have the attachments, I mean, that to me is great" T6 on DiME markerless.*

Technical expertise

All the 6 teachers mentioned this aspect in their comments. They liked the fact that DiME markerless does not require that they learn a new software. *"I would probably get frustrated with the technical part" T4 on DiME marker based". "So I can connect it to my dekstop and connect it to my smartboard and have it up and running in couple of hours." T1 on DiME markerless*

Familiarity

Three of the teachers thought that the students would relate more to the gaming consoles than the marker based systems. They said that the familiarity of the hardware will give them the sense of ownership. T6 said *"The kids are going to relate to that (Kinect and PS Move) because they play those things, they know what those things are. Being that, they have those things at home, just that connection of "oh, I have used that" it's going to have a longer longevity than this that they don't know "oh this is really cool, let me take it apart!" I am just getting you into the mind of an 8 year old".* T1 thought that the gaming consoles would add to their motivation to engage in storytelling. *"You* *are going to become the cool Teacher just because you have an XBox in your classroom. Automatically there will be an interest in this."*

Curriculum match

When asked about whether they would use the system in their classroom, all the teachers suggested that there needs to be a match between what the system provides and what the teacher wants to teach. While the math teacher said that she could not use the storytelling system to teach math in her classroom, two ELA teachers said that they could use it every other day. *"I would probably use it in Reading and writing. I can definitely see this using in Social Studies."* T3 on DiME markerless.

Number of students engaged

Two of the teachers thought the system had limited application in their classes because of the class strength. This could be a future goal for DiME markerless in which the system can track up to 6 actors at a time. But some teachers who had a lesser class strength thought that it was not a problem in their classes. *"If only two kids are gonna do it, what happens to the rest of the class" T5 on DiME marker based*. *"It would be ok if it's not the whole class because when I am doing novels, not all the kids finish at the same time." T3 on DiME markerless*.

5.6. Design guidelines

Based on the findings of the evaluative user study, the design guidelines for the development of motion capture systems for elementary school classroom integration can be summarized as shown in Table 7.

Table 7: Final design guidelines

Constraint	Design Guideline
Administrative	Space
	Cost
	Portability
Technical	Marker placement
	Calibration time
	Technical complexity
Pedagogical	Student engagement
	Curriculum match
	Familiarity

6. DISCUSSION

6.1. Significance of the design guidelines for developers

Interview results from the evaluative user study indicates that the DiME markerless system would fit inside a classroom much better than DiME marker based system. The 9 design guidelines presented in section 5.6 can be applied to any other motion capture system that also requires skeletal tracking and object tracking. While there is some limitations to tracking accuracy and capture volume of the Kinect and PlayStation Move sensors, it could be enough to accomplish learning objectives set by the Teachers. The design guidelines needs to be evaluated further by the developers and researchers in educational technology to test how it might affect their system.

This design guideline was tested in a lab setting on campus because of the time constraints. Ideally, this needs to be evaluated inside an elementary school classroom to get a better understanding. Also repeating this experiment by increasing the number of participants would help in deciding which of the guidelines are more important than others.

6.2. Classroom integration

While many young teachers were excited to use the system in their classrooms, T4 and T5 said that they might not use this system in their classes regularly. This can be attributed to their proficiency and demographic characteristics (years of teaching and age). [11] has shown statistically significant evidence that demographic characteristics will negatively influence their technology integration. T4 summarized this as *"People of my age and my wife's age would take a lot longer to learn the system then some of the younger teachers. They live with this stuff (Kinect and PlayStation 3)"*

Also the DiME marker based system used four tripods to mount 8 cameras with a lot of cables. Teachers suggested that having the cameras mounted on a wall or a ceiling would be a better option in a classroom instead of the tripods.

7. CONCLUSION AND FUTURE WORK

7.1. Conclusion

Motion capture technology is set to hit the field of education in the coming years. There is evidence that the affordances of motion based technology can impact kinesthetic learners. But there is a lack of research about the feasibility of using this technology in classroom environment. This research presented the design criteria that need to be considered while developing motion based applications. An example system DiME markerless was evaluated with a user study with 6 teachers. The results showed that using affordable gaming hardware could work very well compared to expensive marker based setups in classrooms. Statistically significant results were obtained that DiME markerless system was more usable than DiME marker based system. This design criteria could be used by other applications that require motion capture technology integration.

7.2. Future work

Recent advancements in virtual reality has increased the demand for motion capture systems which uses commodity hardware. It will be interesting to test this design guidelines in other virtual reality applications like assistive technology and rehabilitation.

During the interviews the teachers asked about the possibility of adding the following features to help them integrate the system into their classes. Firstly, the ability to track more than just one actor: Since the Kinect can track up to 6 skeletons, this system could be developed further to incorporate this. Secondly, the ability to add backgrounds/characters/objects: One of the teachers thought that the existing system could limit their creativity if the students cannot find the object that they imagined. Hence in the future versions it would be great to have the feature to add custom assets into the system.

It will be beneficial to measure the effectiveness and usage of motion capture in a real classroom settings. This would require collaboration with the School district, teachers and the developers.

REFERENCES

1. Johnson, L., et al., *NMC horizon report: 2014 K*. 2014.

2. Chu, S.L., F. Quek, and K. Sridharamurthy. *Augmenting Children's Creative Self-Efficacy and Performance through Enactment-Based Animated Storytelling*. in *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*. 2015. ACM.

3. Ren, Q., *Applying Motion Capture Technology to Enhance Teaching Quality for Physical Education.* Advances in Information Sciences and Service Sciences, 2013. **5**(3): p. 835.

4. Johnson-Glenberg, M.C., et al., *Collaborative embodied learning in mixed reality motion-capture environments: Two science studies.* Journal of Educational Psychology, 2014. **106**(1): p. 86.

5. Boutsika, E., *Kinect in Education: A Proposal for Children with Autism.* Procedia Computer Science, 2014. **27**: p. 123-129.

6. Giles, J., *Inside the race to hack the Kinect.* New Scientist, 2010. **208**(2789): p. 22-23.

7. Wingrave, C., et al., *The wiimote and beyond: Spatially convenient devices for 3d user interfaces.* Computer Graphics and Applications, IEEE, 2010. **30**(2): p. 71-85.

8. Vrellis, I., A. Moutsioulis, and T. Mikropoulos. *Primary School Students' Attitude towards Gesture Based Interaction: A Comparison between Microsoft*

Kinect and Mouse. in *Advanced Learning Technologies (ICALT), 2014 IEEE 14th International Conference on*. 2014. IEEE.

9. Groff, J. and C. Mouza, *A framework for addressing challenges to classroom technology use.* AACe Journal, 2008. **16**(1): p. 21-46.
10. Wang, Q., *A generic model for guiding the integration of ICT into teaching and learning.* Innovations in Education and Teaching International, 2008. **45**(4): p. 411-419.
11. Inan, F.A. and D.L. Lowther, *Factors affecting technology integration in K-12 classrooms: A path model.* Educational Technology Research and Development, 2010. **58**(2): p. 137-154.
12. Ifenthaler, D. and V. Schweinbenz, *The acceptance of Tablet-PCs in classroom instruction: The teachers' perspectives.* Computers in Human Behavior, 2013. **29**(3): p. 525-534.
13. Inan, F.A. and D.L. Lowther, *Laptops in the K-12 classrooms: Exploring factors impacting instructional use.* Computers & Education, 2010. **55**(3): p. 937-944.
14. Puthenveetil, S.C., et al. *Comparison of Marker-Based and Marker-Less Systems for Low-Cost Human Motion Capture*. in *ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. 2013. American Society of Mechanical Engineers.
15. Chang, C.-Y., et al. *Towards pervasive physical rehabilitation using Microsoft Kinect*. in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference on*. 2012. IEEE.

16. Shotton, J., et al., *Real-time human pose recognition in parts from single depth images.* Communications of the ACM, 2013. **56**(1): p. 116-124.

17. Sanna, A., et al., *A kinect-based interface to animate virtual characters.* Journal on Multimodal User Interfaces, 2013. **7**(4): p. 269-279.

18. Alexiadis, D.S., D. Zarpalas, and P. Daras. *Real-time, realistic full-body 3D reconstruction and texture mapping from multiple Kinects*. in *IVMSP Workshop, 2013 IEEE 11th*. 2013. IEEE.

19. Epstein, Z. *Microsoft says Xbox 360 sales have surpassed 76 million units, Kinect sales top 24 million*. 2013; Available from: http://bgr.com/2013/02/12/microsoft-xbox-360-sales-2013-325481/.

20. Amon, C., F. Fuhrmann, and F. Graf. *Evaluation of the spatial resolution accuracy of the face tracking system for kinect for windows v1 and v2*. in *Proceedings of the 6th Congress of the Alps Adria Acoustics Association*. 2014.

21. Comaniciu, D. and P. Meer, *Mean shift: A robust approach toward feature space analysis.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2002. **24**(5): p. 603-619.

22. Dutta, T., *Evaluation of the Kinect™ sensor for 3-D kinematic measurement in the workplace.* Applied ergonomics, 2012. **43**(4): p. 645-649.

23. Ren, H., W. Liu, and A. Lim, *Marker-based surgical instrument tracking using dual kinect sensors.* Automation Science and Engineering, IEEE Transactions on, 2014. **11**(3): p. 921-924.

24. Han, J., et al., *Enhanced computer vision with microsoft kinect sensor: A review.* Cybernetics, IEEE Transactions on, 2013. **43**(5): p. 1318-1334.

25. Kennewell, S., et al., *Analysing the use of interactive technology to implement interactive teaching.* Journal of Computer Assisted Learning, 2008. **24**(1): p. 61-73.

26. Alibali, M.W. and M.J. Nathan, *Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures.* Journal of the Learning Sciences, 2012. **21**(2): p. 247-286.

27. Cook, S.W., T.K. Yip, and S. Goldin-Meadow, *Gesturing makes memories that last.* Journal of memory and language, 2010. **63**(4): p. 465-475.

28. Chan, J.C., et al., *A virtual reality dance training system using motion capture technology.* Learning Technologies, IEEE Transactions on, 2011. **4**(2): p. 187-195.

29. Yang, Y.-T.C. and W.-C.I. Wu, *Digital storytelling for enhancing student academic achievement, critical thinking, and learning motivation: A year-long experimental study.* Computers & Education, 2012. **59**(2): p. 339-352.

30. Chu, S.L., F. Quek, and J. Tanenbaum. *Performative Authoring: Nurturing Storytelling in Children through Imaginative Enactment*. Interactive Storytelling 2013; 144-155].

31. Chu, S.L., F. Quek, and K. Sridharamurthy. *Ready... action!: a performative authoring system for children to create animated stories*. in *Proceedings of the*

11th Conference on Advances in Computer Entertainment Technology. 2014. ACM.

32. Shiratori, T. and J.K. Hodgins. *Accelerometer-based user interfaces for the control of a physically simulated character*. in *ACM Transactions on Graphics (TOG)*. 2008. ACM.

33. Feltham, J. *Oculus CEO Explains 13+ Age Limit, 'definitely' wants VR for Kids*. 2015; Available from: http://vrfocus.com/archives/16095/oculus-rift-age-limit/.

34. Takala, T.M., P. Rauhamaa, and T. Takala. *Survey of 3DUI applications and development challenges*. in *3D User Interfaces (3DUI), 2012 IEEE Symposium on*. 2012. IEEE.

35. Pearl, T., *Cross-platform tracking of a 6dof motion controller.* Using Computer Vision and Sensor Fusion, Austria, 2012.

36. Greville, T., *The pseudoinverse of a rectangular or singular matrix and its application to the solution of systems of linear equations.* SIAM review, 1959. **1**(1): p. 38-43.

37. Eggert, D.W., A. Lorusso, and R.B. Fisher, *Estimating 3-D rigid body transformations: a comparison of four major algorithms.* Machine Vision and Applications, 1997. **9**(5-6): p. 272-290.

38. Brooke, J., *SUS-A quick and dirty usability scale.* Usability evaluation in industry, 1996. **189**(194): p. 4-7.

39. Bangor, A., P.T. Kortum, and J.T. Miller, *An empirical evaluation of the system usability scale.* Intl. Journal of Human–Computer Interaction, 2008. **24**(6): p. 574-594.

APPENDIX

1. EVALUATIVE USER STUDY PROTOCOL

1. **Demographics(5minutes)**
    a. How long have you been teaching?
    b. What classes have you taught?
2. **Presentation slides about the concept of DiME(5minutes)**
3. **Demo of DiME marker based(10 minutes)**
4. **Qualitative Feedback (15 minutes)**
    a. Could you imagine teaching your class by using this System?
    What did you find confusing in the System?
    b. What did you dislike about the System?
    c. What did you like about the System?
    d. Would it be practical to use this System in School as is? What changes do you think are necessary? Physical changes, Interface changes.
5. **Questionnaire and lesson plan (10 minutes)**

Repeat steps 3,4,5 for DiME markerless (35minutes)

Participant ID : System : DiME marker based

Statement	1	2	3	4	5
1. I THINK THAT I WOULD LIKE TO USE THIS SYSTEM FREQUENTLY					
2. I FOUND THE SYSTEM UNNECESSARILY COMPLEX					
3. I THOUGHT THE SYSTEM WAS EASY TO USE					
4. I THINK THAT I WOULD NEED THE SUPPORT OF A TECHNICAL PERSON TO BE ABLE TO USE THIS SYSTEM					
5. I FOUND THE VARIOUS FUNCTIONS IN THIS SYSTEM WERE WELL INTEGRATED					
6. I THOUGHT THERE WAS TOO MUCH INCONSISTENCY IN THIS SYSTEM					
7. I WOULD IMAGINE THAT MOST PEOPLE WOULD LEARN TO USE THIS SYSTEM VERY QUICKLY					
8. I FOUND THE SYSTEM VERY CUMBERSOME TO USE					
9. I FELT VERY CONFIDENT USING THE SYSTEM					
10. I NEEDED TO LEARN A LOT OF THINGS BEFORE I COULD GET GOING WITH THIS SYSTEM					

Participant ID : System : DiME markerless

Statement	1	2	3	4	5
1. I THINK THAT I WOULD LIKE TO USE THIS SYSTEM FREQUENTLY					
2. I FOUND THE SYSTEM UNNECESSARILY COMPLEX					
3. I THOUGHT THE SYSTEM WAS EASY TO USE					
4. I THINK THAT I WOULD NEED THE SUPPORT OF A TECHNICAL PERSON TO BE ABLE TO USE THIS SYSTEM					
5. I FOUND THE VARIOUS FUNCTIONS IN THIS SYSTEM WERE WELL INTEGRATED					
6. I THOUGHT THERE WAS TOO MUCH INCONSISTENCY IN THIS SYSTEM					
7. I WOULD IMAGINE THAT MOST PEOPLE WOULD LEARN TO USE THIS SYSTEM VERY QUICKLY					
8. I FOUND THE SYSTEM VERY CUMBERSOME TO USE					
9. I FELT VERY CONFIDENT USING THE SYSTEM					
10. I NEEDED TO LEARN A LOT OF THINGS BEFORE I COULD GET GOING WITH THIS SYSTEM					

2. FORMATIVE USER STUDY PROTOCOL

1. **Demographics**

How long have you been teaching?

What classes have you taught?

2. **Technologies(Hardware/Software) used during teaching**

What software/hardware do you use in your teaching?

What happens if a device/software isn't working correctly?

Is there a computer and a large screen display in every classroom?

How do you manage situations when there are more students than Computers?

What are some of the challenges of using these Technologies for teaching?

Do they prefer using applications over the Internet, Desktop or Tablets?

Specific examples.

How much Technology budget does your School have? What are some of the existing applications you use?

3. **Presentation of DIME**

Powerpoint presentation

4. **Classroom integration**

How much Enactment space is typically available in a classroom?

Would it be practical to use this System in School as is? What changes do you think are necessary? Physical changes, Interface changes.

What classes could benefit from a tool like this?

If you had this System at school, how would you conduct classes using it?

Describe where this system could be placed in your school/classroom?

How would you involve an entire class to use this System?

5. **Feedback**

Could you imagine teaching your class by using this System?

How useful did you think the objects in this System were?

How could children make the best out of this System?

What did you find confusing in the System?

What did you dislike about the System?

What did you like about the System?

What approval would you need to use this in your classroom?