An Immersive Motion Capture Environment

Daniel Kade, Oğuzhan Özcan, and Rikard Lindell

Abstract—Motion capturing technology has been used for quite a while and several research has been done within this area. Nevertheless, we discovered open issues within current motion capturing environments. In this paper we provide a state-of-the-art overview of the addressed research areas and show issues with current motion capturing environments. Observations, interviews and questionnaires have been used to reveal the challenges actors are currently facing in a motion capturing environment. Furthermore, the idea to create a more immersive motion capturing environment to improve the acting performances and motion capturing outcomes as a potential solution is introduced. It is hereby the goal to explain the found open issues and the developed ideas which shall serve for further research as a basis. Moreover, a methodology to address the interaction and systems design issues is proposed. A future outcome could be that motion capture actors are able to perform more naturally, especially if using a non-body-worn solution.

Keywords—Immersive acting environment, Interaction in a mediated environment, Motion capturing, MoCap.

I. INTRODUCTION

Motion capturing (MoCap) has already been used for quite some time in the movie and games industry [1], [2]. For these kinds of applications, actors, stuntmen and athletes are used to record motions and to bring virtual characters to life through acting performances made closer to reality. Through observations and interviews, explained in more detail below, we came to the conclusion that the current acting environment in a motion capture studio does currently not provide a natural acting environment for an actor if we compare to a stage or film acting environment. This is because the virtual environment an actor is acting for is not visible and perceivable for the actor while he is acting. Objects, obstacles, other virtual personas or events have to be memorized and are not yet visualized in real-time without having the actor to turn his head towards a screen while acting. In a traditional acting environment, an actor is surrounded by a suitable environment, proper physical objects and real personas which allow the actor to concentrate on the acting itself. For current motion capture shoots, an actor has to memorize the scenario and do the acting. The performance is very much dependent on the actors capability to imagine the scenery and to put himself in the desired role and mood. This is why, it is in many cases more challenging to attain the same level of emotion and naturalness possible in more immersive types of acting can seem less natural. Moreover, the acting outcomes are in many cases less accurate and recorded performances can be less believable. This also effects the emotional connection of the game player in the final application. The player can see and feel when movements and emotions are not close to realism. A common practice to overcome the mentioned issues is to correct, adjust or augment the recorded MoCap data through post-production processes and to improve the results through animation.

In this paper we provide a state-of-the-art overview of the related research areas and address current issues for actors in a motion capturing environment, as well as current issues in terms of technology and systems design to create a more immersive motion capturing environment. It is hereby the goal to make aware of the found open issues for further research. Research through observations, interviews and questionnaires has also been conducted to reveal the challenges actors are currently facing in a motion capturing environment. Moreover, we are opening possible design and technology solutions to reduce or eliminate the addressed problems for discussion. A methodology to address the interaction and systems design issues is proposed, as well as a proposal on how those issues are intended to be solved in a future research.

II. CURRENT STATE OF THE ART

Solving the above-mentioned problems implies three different research areas. The first research area considers the visualization of virtual content while acting. The second research area needs to consider motion capture acquisition, tracking and the orientation of multiple actors in the motion capturing shoot volume. For the last research area, the interactions of the users inside and outside the scenario as well as the design of the user interface needs to be considered. The parts, which are of main interest to create a more immersive acting environment, are the creation and visualization of a virtual and immersive environment, as well as the interactions with and within the environment. Each of the mentioned research areas have possible solutions to partly solve the addressed problem. An overview is provided below.

A. Visualization

The option that, when improved, should yield the greatest positive result in actor performance is to make a motion capturing environment more immersive by making the virtual content visual. Several applications that offer this ability are already available:

A mixed reality environment, which could possibly be used in a motion capturing application, is used for many industries such as military-based training environments. In this respect, some research projects are capable of providing an immersive environment without using augmented reality (AR) glasses [3]. To achieve this aim, an immersive environment is created.

D.Kade and R.Lindell are with the Department of Innovation, Design and Engineering, Mälardalen University, Västerås, Sweden (e-mail: daniel.kade@mdh.se, rikard.lindell@mdh.se).

O. Özcan is with the Department of Media and Visual Arts, Koç University, Istanbul, Turkey e-mail: (oozcan@edu.ku.tr).
by using a mixture of real and projected objects, transparent digital flat screens and the capability to add smell and temperature changes to the environment. There is also a wide-area mixed reality application, which was realized to create an immersive virtual reality environment in which users can walk and run freely among simulated rooms, buildings and streets. In such applications, large rear-projection screens that employ digital graphics are used to depict a room’s interior, a view to an outside world or a building’s exterior. The applications also provide life-size projection displays with physical props and real-time 3D graphics [3]. A scenario like mentioned above has its limits when considering it for a motion capturing shoot. Large projection walls or other props cannot be placed in front of the optical-based motion capture cameras because the recordings would be blocked and occlusions would occur. Setting up such a training environment requires time and planning which is not economical in a MoCap environment because the scenery for a motion capture shoot changes often and needs to be dynamic. Therefore, this solution does not offer the right technologies to be used in general to create an immersive environment for MoCap actors to support more natural acting.

A novel optical see-through head-worn display that is capable of mutual occlusions could also be considered for motion capture shoots. Here, mutual occlusion is an attribute of an augmented reality display where real objects can occlude virtual objects and virtual objects can occlude real objects [4]. Mutual occlusions are one of the problems in visualizing 3D content in a real world environment. Research is also conducted to test the perception of image motion during head movement [5]. The perception of image and head motion is tested when wearing a head-mounted display (HMD). In another significant work, in addition to virtual reality (VR) and AR, two control conditions were studied: viewing real-world objects, and viewing real-world objects through a head-mounted display. The presence and absence of motion parallax was crossed with all conditions. Like many previous studies, another research found that depth perception is underestimated in virtual reality, although the magnitude of the effect was surprisingly low. The most interesting finding is that no underestimation is observed in AR [6]. Therefore, the VR technology is not so suitable to create a natural, immersive motion capturing environment for actors. Another issue is that for motion capturing shoots, the interactions between actors and the possibility to see props are important. This is another reason why a VR environment is not the best solution when an actor needs to be able to see real world objects and persons, as well as virtual content at the same time. A general problem with HMD’s remains, regardless of future AR glasses: a significant part of the actor’s face is covered and can limit MoCap shoots. The glasses especially limit the freedom of movement of an actor and the shoots are limited when facial captures are of importance.

To make virtual environments as visually immersive as possible another solution has been widely researched by using flat panel 3D displays [7]. Equipping an entire motion capturing room with these 3D displays, or even by just using a single flat screen needs to be investigated if this might be a solution that could create an immersive acting environment.

Another method of possible use to display virtual content to the actor while acting is by using the emerging laser based pico projector technology [8]. These projectors are of small size [9], [7] and come in different technologies based on micro-LCDs, the Texas Instruments’s DLP technology, which uses an array of microelectromechanical systems (MEMS) micro-mirrors and LEDs, or projectors based on laser scanning [10]. The projected data can be shown on small screens or can be reflected to polarized video contact lenses which a user is wearing [7]. A screen, which is placed right in front of the actor might limit the actor in some cases but might be applicable to some MoCap shoots. Nevertheless, it would be good to answer the question on how virtual content can be shown to the MoCap actors and it needs to be researched how new technologies can be applied to a MoCap scenario.

B. Tracking

The tracking of a person’s eye, head and body position is an important factor to create an immersive virtual environment. For many visual solutions it is important to determine the exact eye, head and body position to create an immersive and accurate virtual environment. In the tracking research area, there are multiple usable approaches to help developing an immersive MoCap environment. One approach is presented as a computer vision based head tracking system for augmented reality. A camera is attached to a head mounted display and used to track markers in the user’s field of view. Another tracker is attached to the HMD and can be tracked by an independent camera, mounted on the ceiling [11]. Also, in the research area of eye tracking there are some different approaches, which might be of use to create an immersive acting environment. For example, the methods and technical challenges of low-cost eye tracking, as well as the design decisions were discussed to integrate eye tracking to everyday human computer interfaces [12]. Tracking solutions for head tracking get more advanced. It is described how a new head tracking method utilizes a gyroscope mounted on a head-mounted display and a bird-eye view camera that observes the HMD from a fixed third-person’s viewpoint. Furthermore, an extension of this method to a hybrid registration, combining it with a user’s view camera is proposed [13]. However, a solution to suit the needs of a dynamic system of body- and non-body-worn technology needs to be found to solve the tracking of multiple persons and their eye and head orientation at the same time.

C. Interaction

We consider designing useful interaction scenarios for an immersive MoCap system as important and believe that research on the interaction and perception in virtual environments could deliver important insights to interact within an immersive motion capturing system. In this respect, some research, which works on the body-centered interaction in immersive virtual environments, provides such insights and shows how interaction and navigation in such virtual
environments might be possible [14]. There is also some other research on how it is possible to approach the problem of making virtual walking more real. In this research, the problem of having a large virtual space but just a limited physical space is addressed by introducing a treadmill control algorithm to solve this problem [15]. Some research has been conducted on solving the issue of virtual walking by using treadmills and treadmill algorithms [16]. This body of research which was done for other purposes might be adaptable to develop more immersive environments for game-based motion capture shoots; but it still needs to be answered, how more natural actors performances during motion capture shoots can be achieved by using these techniques. As well as how this might work for multiple performers.

Motion capture techniques were also adapted to an interactive dancing game by another research project [17]. This kind of research, aims at developing a system to implement a visual interaction between a human and a virtual avatar through the use of motion capturing technology. In such a system, a virtual partner is able to recognize and respond to the player’s movements without a noticeable delay. The completion progress of a move is tracked progressively and the virtual partner’s move is rendered in synchronization with the player’s current action. The virtual dance partner is displayed on a large screen and the motions are captured with a standard marker-based system [17]. The set-up developed in this research does allow the player to adapt his motions to the displayed content but it limits the player by always having to look at one fixed screen. This means that the flexibility and the naturalness of movements get lost. This is an issue that needs to be solved for a MoCap set-up where precise movements have to be performed and captured.

Research that combines AR and gesture-driven interfaces has been conducted by creating a system that uses body motion to visualize and interact with virtual objects populating AR settings. Body motion is used to implement a whole body gesture-driven interface to manipulate the virtual objects. Gestures are mapped to correspondent behaviors for virtual objects, such as controlling the playback and volume of virtual audio players or displaying a virtual object’s meta data [18]. Gestures that might be used in a system that supports the acting in a MoCap environment should not be predefined like in the above-mentioned research. This means the actor needs to have the freedom to interact with objects in a, for him, natural way. Up to now this is still an issue, which would be good to be solved.

In a virtual environment the action and reaction of virtual personas and objects is in most cases fairly primitive. In some research characters in virtual environments have been trained to react to predicted and unpredicted events in order to maintain realism. The approach to solve the problem of motion synthesis for interactive, humanlike characters was by combining dynamic simulations and human motion capture data [19]. It needs to be investigated, how and if the virtual content in a MoCap environment can be used to steer the actors while acting. Other than in the above-mentioned research the actors do not directly steer or interact with virtual content, they far more use virtual and mediated objects like real objects.

D. Combined Research Areas

Research, which combines one or more research topics like visualization, motion capture, tracking and interaction could be partly or fully considered. The combined research for actors who should perform effectively and naturally for motion capture (MoCap) shoots. There are several approaches to be considered when creating a scenario as close to reality as possible:

Using the sense of touch is one way of raising the perception in such an environment. In this respect, some research focuses on a system that answers the question on how to substitute a persons body in a virtual reality by a virtual body or avatar. In this kind of research, the avatar is seen from a first person perspective, moves as the person moves and the system generates a touch feeling on the real persons body when the avatar is touched. Such a replacement of the persons real body by a virtual body requires a wide field-of-view head-mounted display, real-time whole body tracking, and tactile feedback. In such research, a tactile vest is used to trigger the haptic feedback of touch and to raise the immersive experience [20]. In a similar approach, the sense of touch in an augmented reality environment is introduced by using a haptic glove. This allows an AR system where people can see a life-size virtual human avatar in a real environment and are able to have a physical haptic interaction with the avatar [21]. This kind of research aims at working on displaying and controlling an avatar and getting haptic feedback when the avatar is touched. For MoCap, haptic feedback could be an important aspect to consider when the virtual environment shall be more naturally perceivable. Nevertheless, there are indications that actors need to wear some equipment to create an immersive environment which could limit the actors in their movement and acting. Motion capturing techniques are essential to create realistic character movements for video games or animated movies. However, these techniques are also used in other kinds of applications and research. One of them is developed to understand how a real-time motion capture technology can be used on a live theatrical performance. In this research, the real-time motion capture data is streamed on a multi-screen topology, while real-time generated virtual scenery is creating the virtual context of each act as a background. On-stage actors are interacting with digital avatars that are controlled by actors wearing a motion capture suit throughout the theatrical act. A screen, which is installed in front of the on-stage actors, displayed the avatars controlled by the motion capture (MoCap) actors who performed their acting in real-time on a close-by motion capturing area [22]. In the mentioned research, the displayed scenery and avatar content around the actors was used to steer and drive their acting. How this can be applied to an immersive MoCap environment in terms of technology and interaction still needs to be solved.

III. CONDUCTED RESEARCH

The above-mentioned technologies and research projects
show a potential use to create an immersive MoCap acting environment. Nevertheless, less research has been done on using the mentioned technologies in a motion capturing environment and for motion capturing actors. Therefore, it was of importance to identify the challenges that actors are facing in a current motion capturing environment before designing a solution. For this matter, we used direct observation as a qualitative method to observe two motion capture shoots (3 actors) and one audition (24 actors), which have been observed in a professional motion capture studio. Moreover, interviews with directors and technicians at the MoCap studio have been performed and questionnaires have been handed out to experienced MoCap actors.

One MoCap shoot was done by a performer that has never done acting or MoCap shoots before but was familiar with the environment. Within this 4 hour shoot, simple full-body movements for a game were captured. Mostly walk, run and idle movements with and without weapons needed to be performed. In the second MoCap shoot that has been observed, two experienced MoCap actors performed in-game cut scenes and in-game character movements. The recordings for this shoot took 8 hours. One actor has worked for 8 years in motion capturing and is a trained stuntman with some acting experiences. The second actor was also a stuntman that has worked for 4 years with motion capturing. After the shoot a questionnaire was handed out to the two actors. The focus of the questionnaire was on identifying challenges of the acting in a virtual environment and on identifying the feeling of immersion in a current MoCap shoot.

For the audition that was also observed, 24 actors with different backgrounds in drama, film, commercials and stage acting have been invited. Two of the 24 actors have done motion capture shoots before. Every actor needed to perform the same tasks which consisted of movements, short acting performances for game scenarios and a monologue. For the audition each actor had about 15-20 minutes to perform. The main goal for all observations was to see what challenges experienced and inexperienced actors are facing in a MoCap environment and to develop ideas on how to overcome these issues.

In addition, several interviews with directors and technicians have been done to discuss the challenges actors are facing during a MoCap shoot and also where the technology limitations for MoCap shoots are at the moment, especially when it comes to the practical use.

IV. FINDINGS

The outcomes of the observations and questionnaires are described below and have shown expected but also unexpected results. For the evaluation of the collected data it is essential that we consider the difference in the observed user groups. Therefore, the findings are split into ‘Inexperienced MoCap Actors’ and ‘Experienced MoCap Actors’.

Inexperienced MoCap Actors: All actors observed at the audition as well as the inexperienced actor at the MoCap shoot belong to this group because most have never done a MoCap shoot or even been in a MoCap studio. Although 2 actors which were at the audition participated at a motion capture shoot before but only once or twice. This is why they are also considered as inexperienced MoCap actors. Many actors that belong to this group had a similar reaction while they were performing their prepared tasks in this, for them, new environment. Similar question like: “Is there no one to act to?” were asked throughout the day. Some actors were even surprised that they are standing in the middle of an empty room with only some minor props like wooden boxes or weapons that were at their disposal. One actor even said during his performance: ”It is weird that there is nothing”. This was one of the reactions we expected to see because we assumed that the different way to work and the environment which needs to be imagined, is new to most actors. One task that the actors at the audition had to prepare was to play the role of a spy that needs to break into a facility which is protected by a laser alarm system. Many details to describe the scene better were purposefully left vague in order to see how well actors coped with an environment where much is left to imagination. The director mentioned that this technique is a recognized factor of MoCap acting so that it was amplified for the audition process. Observing the actors performance of this task revealed another issue many inexperienced actors had while performing for a MoCap shoot. It seemed to be hard for many actors to imagine where virtual content and partners in an empty room are, even though the scene was described in more detail. Even more important it was observed that many hand and body movements were lacking in correct and accurate motions. This resulted in many cases in movements that did not look natural and believable. For example, many actors had their issues was to avoid the laser alarm system. Actors imagined a laser grid in front of them and acted through an acrobatic solution to avoid the beams. The issue was very often that an imagined laser beam was very clearly avoided with one foot or arm but then later crossed by a different arm or leg. Another example of the same issue was when actors imagined use of an electronic devices like a laptop or door security systems. Their movements, while simple where not believable and not accurate enough. It seemed that the actors had issues to perform an accurate motion sequence to avoid the imaginary laser grid out of their memory. It was fairly unexpected that most inexperienced MoCap actors had these problems. Other tasks that actors were used to from their normal acting environment or acting education seemed to pose for the majority of the invited actors no problem. Especially when they had to perform a monologue or play a role where the acting is more important than the motions and the use of the virtual environment. To conclude the observation of the actors at the audition, the performances of the actors were very much dependent on their imagination and skills to bring in naturalness and play a character believable. Nevertheless, most inexperienced MoCap actors do have some issues to perform sufficient enough in this rather unknown work environment without a prior training. For the actor who did the professional MoCap shoot, no issues were identified. This was rather unexpected.
but might have been because he was used to the environment, even though he is not a trained actor. Also the task was simply to perform movements without acting that involved common movements that do not need special training like running, walking, jogging with and without a weapon.

Experienced MoCap Actors: The observation of the experienced MoCap actors has shown that the issues inexperienced MoCap actors had were not identifiable. What the actors did to make movements more believable and appealing for the virtual character was to adapt their motions and to repeat scenes multiple times to improve the naturalness and believability of the character they play. For this it was important to get themselves quickly into different moods and emotions. This usually happened by shouting, performing several motions before the recording or by hitting themselves, for example to get in an aggressive mood. An issue that occurred throughout the shoot day is that virtual content was simply overseen or forgotten. In many cases this was overcome by using props but in some cases it is not possible to put props in the way of an actor because of safety reasons when the actor needs to move in a more acrobatic way. The questionnaires that were handed out to the experienced MoCap actors revealed that one of the challenges the actors described is finding the right mix between what is realistic and what is believable when the motions are used for a game. Furthermore, the actors added that for a MoCap shoot you cannot cheat because the MoCap camera capture your motions from every angle. A major conclusion we identified after all observations and interviews with directors of the MoCap studio is that the performance is very much dependent on the actor capability to imagine the scenery and to put himself in the desired role and mood. This is why, in some cases the captured performances of a motion capturing shoot appear less emotional and motions seem less natural. Moreover, the acting outcomes are in many cases less accurate and recorded performances can be less believable. Furthermore, it is also helpful to know what is important for a MoCap actor to deliver a good performance.

V. OPEN ISSUES

Addressing the issues to create a more immersive acting environment for motion capturing is one of the main goals of this paper. This section lists and describes the issues identified through the conducted research and the literature review. One identified issue is the fact that motion capture actors cannot yet see and especially feel the virtual environment they are acting in while they are acting. This needs to be solved to allow actors to perform in virtual worlds easier and more naturally. To create an immersive motion capture acting environment, research needs to be done to find the best technology solutions, which are of use in a MoCap studio. Therefore, it needs to be considered that the goal of a motion capturing shoot, to capture the motions of actors, is not affected by the system to be developed. For example, the recording quality of the motion capturing cameras should not be affected and the vision to markers that shall be tracked should not be blocked. Another point the needs to be considered when designing an immersive acting environment is that the actors motions and emotions are not limited or affected by such a system. For example, an actor should not need to unintentionally turn his head to be able to see virtual content, otherwise these motions are captured and the end result looks less believable. Interactions and the design of user interfaces for an immersive MoCap environment are, as far as the literature review has shown, fairly unexplored. It would be good to answer the question on how interactions with and within such an environment are possible. We believe that this could help the users of such a system to interact with virtual content and to create a more dynamic acting experience. In current MoCap production processes, captured inaccurate and unnatural motions and emotions, as well as smaller issues that occurred while recording like uncaptured, occluded or flipping markers are corrected through post-production processes and animation. Depending on the quality of the MoCap shoot and the performance of the actors the effort for post-production and animation varies. Therefore, it would be good to address this problem and we believe by solving the above-mentioned issues, the time for post-production and animation can be decreased.

VI. FUTURE SOLUTION

It is intended to solve some of the addressed problems by conducting research and implementing a prototype. The main focus hereby will be set on designing and applying techniques and technologies to a MoCap environment to create an immersive MoCap environment. Furthermore, a focus will be set on how interactions with and within such an environment are possible and how the environment improves a MoCap shoot. For this it is intended to do more tests and interviews with more experienced MoCap actors.

Parts of already conducted research in the areas of AR and VR visualization, tracking and interaction in an immersive virtual environments might be considered to create the system. Of particular importance, the advancements in visualization technologies are an important aspect to create a more realistic scenario. Therefore, a suitable solution for a MoCap environment needs to be found by considering the problems addressed in this paper.

Towards this future solution it is intended to combine the existing technology at a motion capturing studio with the immersive virtual acting environment. Combining different technologies will be an essential task. Especially, the visualization technology needs to be combined with a tracking system that considers head, eye and body positions of multiple actors to create a useful immersive visualization. Nevertheless, an actor should be able to perform his acting, in terms of body- movements as naturally as possible. This means that solutions to solve the mentioned problems should be as non-body-worn as possible. When combining the existing MoCap system with the immersive virtual environment it is also intended to allow interactions between the actor and virtual content, as well as to allow interactions between the technicians and the system. Therefore, guidelines or further research outcomes in interaction, navigation and design
guidelines for virtual environments could serve as a basis for a way of interacting in a motion capture environment. For our future solution we intend to use a methodology that we are developing to address the identified issues through an approach that combines interaction design research methodologies with computer science methodologies. First a concept creation phase that involves participatory design methods (PDM) such as learning-by-doing, co-creation and low-fidelity mock-ups. The personas of the environment will be defined and their needs will be considered from the beginning of the research. An extensive reality check to secure the set goals and their reachability. Afterwards, a conceptual prototype will be implemented by using common interaction design and computer science research methodologies, such as workshops, iterative design cycles and agile programming methods. A final working prototype, as well as interaction design guidelines shall show the results as well as the found solutions to the addressed problems.

VII. CONCLUSION

From our observations we gained the knowledge that a performance is very much dependent on the actors capability to imagine the scenery and to put himself in the desired role and mood. Also, the results of a motion capture shoot and the amount of post-processing work is dependent on the quality of the actors’ performances. Throughout the observation of inexperienced MoCap actors, many actors were surprised that they were standing in the middle of an empty room and reactions like: "It is weird that there is nothing" and "Is there no one to act to?" were mentioned. Moreover, a given task revealed that imagining virtual content and their positions to perform accurate motions posed problems to some inexperienced MoCap actors. We also observed that it was challenging for some inexperienced motion capturing actors to adapt to the acting environment quickly. In many cases the performances of inexperienced MoCap actors appeared less emotional and less natural. For experienced MoCap actors the major issue observed, was that motions needed to be repeated in some cases multiple times to meet the directors expectation for the character being portrayed. Another reaction was that it took some time for the actors to get in the right mood, which they sometimes needed to change quickly between different takes. Therefore, actors used specific movement sequences, sentences, were hitting themselves or were screaming to get in the right moods and motions of the character. The questionnaires, the experienced MoCap actors filled out, revealed that in the actors’ opinions it is challenging to find the right motions that suit the character to be played and to perform believable movements for games.

As there is no solution yet solving the issues of acting in a visually and emotionally perceivable acting environment using virtual environments, we believe a successful solution will improve not just the actual performances and the recorded MoCap results, but also decrease the post-processing time and minimize the number of corrections recorded to make up for imprecise movements. It is also believed that a solution to the above-mentioned problems allow motion capturing actors to deliver a more natural and believable performance. Such an environment could also help the actors to get in the right mood and role of the character to be played easier.

By addressing the open issues of a visually and emotionally perceivable MoCap acting environment, we opened the discussions to solve the related problems. In the future we will be working on creating an environment which is visually and emotionally perceivable and uses as little body-worn equipment as possible. We also intend to research on how interactions within an immersive MoCap environment can be made possible and what tools could support the interactions.

REFERENCES


