Augmented Virtuality in Real Time for Pre-visualization in Film

Alex Stamm* — Purdue University

Patrick Teall[ŧ] — Purdue University

Guillermo Blanco Benedicto[ŧŧ] — Purdue University

Abstract

This project looks into creating an augmented virtuality pre-visualization system to empower indie filmmakers during the on-set production process. Indie directors are currently unable to pre-visualize their virtual set without the funds to pay for a high-fidelity 3D visualization system. Our team has created a pre-visualization prototype that allows independent filmmakers to perform augmented virtuality by placing actors into a computer-generated 3D environment for the purposes of virtual production. After performing our preliminary usability research, we have determined a clear and effective 3D interface for film directors to use during the production process. The implication for this research sets the groundwork for building a pre-visualization system for on-set production that satisfies independent and emerging filmmakers.

Keywords: indie filmmaker, pre-visualization, virtual production, augmented virtuality, compositing, chroma-keying

Figure 1: Actor without pre-visualization in a green room.

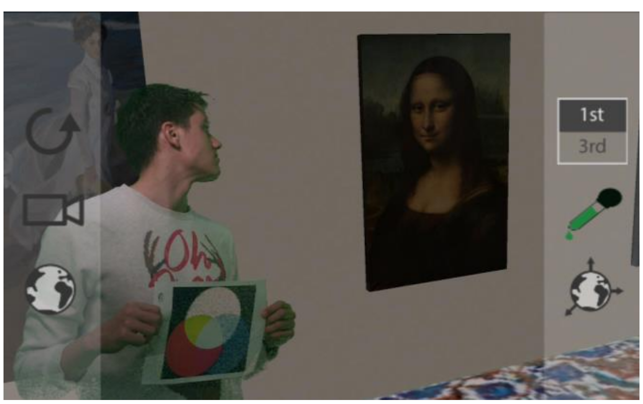

Figure 2: Actor with pre-visualization in a virtual museum.

* stamm@purdue.edu

ŧ p.teall25@gmail.com

ŧŧ gblancobene@gmail.com

1 Introduction

Imagine for a moment the scenario of a director on-set in a green screen room. The director places the actors, sets up the cameras, and the shooting process begins. The film about to be shot will include high-fidelity special effects and the scene will have to be recreated in the post-production process. However these visual additions are created later on in the production by visual effects specialists. Because the filming process must take place first in a green screen room, the problem lies in that the director must imagine this scene in his or her head, including the placement of the actors and what the scene might actually look like later on.

A pre-visualization tool would give this director a clearer picture of what the scene might look like, decrease the risks that come if it were to look different than imagined, and enhance the production timeline as more people would have the ability to see the same visualization. Communication would be greater among the production team, preventing errors that might occur later on in the production pipeline.

What our team is investigating is the possibility of this kind of pre-visualization being done on a mobile platform, with a small hardware footprint, and usable in the on-set production environment. Current systems require expensive hardware and large camera positional-tracking systems that must be built into the green screen room itself. Because of the hefty price tag of these systems, they are not available to the low-budget filmmaker who would like to use special effects in his or her production. To our knowledge, there is no alternative for these indie filmmakers to use pre-visualization in their films due to the costly nature of these current systems.

To investigate this problem, our team is leveraging technology of Google's Project Tango development tablet to build an inexpensive and mobile application alternative to current on-set pre-visualization systems for film. The Project Tango combines 3D motion tracking with depth sensing to give the device the ability to know its location and how it is moving through space, giving a correlation between a real camera and a virtual scene camera. By utilizing real-time rendering, our application gives a live preview of what a final rendering could look like from a preliminary perspective. In addition, using real-time rendering for our pre-visualization system allows for user in the planning of complex motion for shots, giving the filmmaker important information about the scene before the actual shooting takes place. [2] Performing this pre-visualization step prevents errors and reshoots, subsequently saving money for the production and enhancing the film itself as these savings can be allocated to other areas of the production.

2 Background

Pre-visualization as a practice has been used since the beginning of film production in the form of storyboarding and animatics to help the director and production team understand the perspective before the actual shooting takes place. [5] Having this information organized before the onset production process is vital to the film production. The intention for this practice is to minimize errors and reduce the need to reshoot scenes later on in the production.

Pre-visualization is necessary for the film industry in using virtual assets for its productions. Directors must understand where actors are placed in a virtual scene, if they are interacting with 3D animations, and where visual effects need to be placed in the scene. Because a film set is a fast-paced environment where time and money are limited resources, pre-visualization must be performed quickly and is used for the purpose of framing, movement, blocking shots, and staging the scene; sacrificing quality, in favor of speed and precision. [5]

The role of an on-set pre-visualization system is to achieve the insertion of computer-generated assets into video. Traditional methods of compositing these elements into the scene involve the placement of 2D elements alongside the actor to get an idea of how the scene might look with all assets combined. However, these compositing techniques are limited to pre-placement of the physical overlay, with the scene being interrupted as a 2D medium. [2]

Tracking techniques used in films involve the placement of markers in the green screen that tell the camera the 3D perspective of the room. When tracking techniques work, the camera can export this perspective information into post-production compositing software that correlates the position of the virtual scene with the perspective of the actual film camera. For on-set pre-visualization, the position of the camera needs to be estimated in real-time in order to perform this perspective tracking of the virtual scene. The use of this "pose estimation" for the camera is required for pre-visualization, so that it correlates with the real film camera as the filmmaker moves the camera around a production set. [2]

Predecessors have described this system as "augmented virtuality." This mixture of real objects and virtual environments is located on the mixed reality-virtuality continuum. [4] This continuum describes visual displays on a spectrum with the far left being real environments and the far right being a total virtual environment. Augmented virtuality exists within this spectrum, closer to the side of a complete virtual environment. Where augmented reality is having virtual objects interact with a real environment, closer to the left on the spectrum, augmented virtuality is having real objects interact with a virtual environment, closer to the right on the spectrum. Describing our visual display system in these terms helps determine the level of mixed-reality that we are working with in our system.

In order to assess the quality of a virtual reality application, systems like the VRUSE have been put in place to measure the general usability of virtual reality applications. The VRUSE is a special questionnaire designed to measure the usability of a virtual reality system. It is used as a computerized diagnostic tool for evaluating virtual environments and the user interface required to navigate through these environments. This questionnaire is a compilation of 100 questions with each question in the survey pertaining to both a usability and diagnostic factor in the user interface. This questionnaire is designed according to established statistical principles on a five-point Likert attitude scale ranging from 5: strongly agree, to 1: strongly disagree. The user responses dictate how to diagnose the system. The measurement of usability include factors such as functionality, consistency, and immersion, and the diagnostic factors include metrics such as ease of use, learnability, and intuitiveness. [3]

3 Methodology

To build this platform for the purposes of pre-visualization, we designed our platform with the hardware capabilities of Google's Project Tango inside the environment of the Unity game engine. Google has developed a Unity Software Developer's Kit so that the 3D camera and virtual scene construction capabilities of the Project Tango can be utilized in Unity's game engine editor. [6] This functionality can only be used once the application is built to the Project Tango Android tablet, which is the device responsible for conducting our usability testing.

To remove the background of the green screen, we used a well known technique called chroma-keying. The chroma-keying functionality of our prototype is based on basic pixel subtraction from a custom-made shader algorithm. This custom shader limits the user in their ability to refine the matte for proper color correction and in-depth compositing since it provides the basic pixel subtraction functionality for our user.

After developing this platform, we then gathered participants for the purposes of testing. These participants all have a strong background in film and experience on a live film set. We tested these participants by performing a contextual inquiry to understand how these participants interact with the device and how they perform chroma-keying on a real green screen set. After each contextual inquiry, we performed an ethnographic interview to understand the thought process of each participant immediately following using the application. Here we learned how these participants might use a system like this in the future and were able to receive immediate feedback about their thoughts on the system. Immediately following this interview, we gave our users the customized VRUSE questionnaire which gave us more quantitative information about the usability of the system. In addition, this questionnaire also held open-ended questions so we could receive qualitative information about the usability of the system.

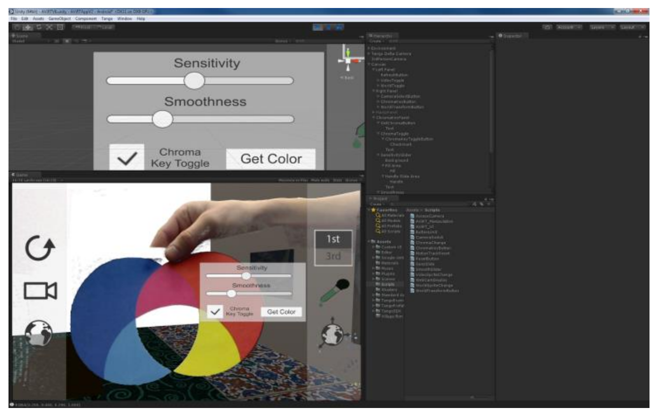

Figure 3: Unity game engine development environment.

3.1 Participants

We assumed that our research participants are not familiar with using virtual assets in their film productions. The participants used are people who are familiar with using mobile applications, and our subjects are people who are involved with film studies and have experience participating in the film production process.

During our usability research, we restricted the user to a single pre-set 3D virtual environment. We limited our users to one-click interaction to give them the power to incrementally manipulate the virtual environment.

3.2 Cognitive Walkthrough

Before beginning the development of our application, we asked an industry independent filmmaker about his understanding of virtual assets, green screens, and 3D environments. This step was important for our team of computer graphics students to step into the shoes of traditional filmmakers that are unfamiliar with visual effects productions. By doing this, we were able to identify what interaction would be necessary for a director to translate their vision for the film into a digital rendering in a short amount of time. Through this cognitive walkthrough, we identified flaws in our initial design and some of the potential usability problems.

3.3 Contextual Inquiry

After the application was developed, we tested it with a group of students from the film and video studies department at Purdue University and the telecommunications department at Ball State University. For this study, our sample consisted of students, faculty, and alumni from these departments in order to represent various skill levels of student and independent filmmakers. The usability testing that we conducted was in the form of a task-based contextual inquiry to give us insight into the learnability of our application, the effectiveness of the tool on a film set, and the immersion into the virtual scene. For this purpose, we provided participants with a standardized sample scene, the Tango device with our application, and a green room film set. The metrics evaluated included task success, errors, efficiency, and learnability. All tests included an actor, conductor, note-taking observer, and observational video recorder. The Tango screen, and the entire room, were recorded to give insight into the overall movement and direction around the set. Participants were debriefed with the purpose of our project, and they were asked to verbalize their thoughts for the duration of the tasks. The scene that was used for this testing was a 3D museum filled with paintings. Participants were asked to place the provided actor into the museum as if they were actually observing the Mona Lisa and Persistence of Memory paintings as realistically as possible.

3.4 Ethnographic Interview

A multitude of questions were asked throughout the contextual inquiry and following the completion of the tasks. This provided a chance for informal discussion with the participants giving insight into the needs of the industry and their opinions of the application itself. Here we were able to identify the user pain points and frustrations associated with the application. Pain points included difficulty with manipulating the virtual set and inability to align actors with the ground of the virtual set. The dialogue provided by participants affirmed that a compass alignment for manipulation would be a necessary update and that swipe manipulation would be intuitive but inappropriate for precise alignment. Participants also mentioned that tilting the virtual set may successfully 'ground' actors in the scene. Most responses were positive, with claims that the application was extremely easy to use, learnable with little to no direction and was valuable to them, even in its first iteration.

Figure 4: Contextual Inquiry walkthrough at Purdue University.

3.5 Post Evaluation Questionnaire

Our post evaluation questionnaire was a modified version of the VRUSE usability questionnaire hosted by Qualtrics. This questionnaire is a computerized diagnostic tool for the usability evaluation of virtual/synthetic environment systems. [2] The full questionnaire consists of 100 questions that combine a list of usability and diagnostic factors for determining the total usability of a synthetic environment. The questionnaire is broken into ten usability factor sections that can help diagnose the system. After consulting with the Department of Statistics at Purdue University, we have developed a modified version of this questionnaire, reducing the number of questions from 100 to 69. The purpose of this reduction was to remove sections that did not adhere to the scope of our project. One reduction includes removing the user guidance and help section which is tailored to the user being able to request help for online assistance.

We extended our standardized questionnaire by adding open ended questions to allow participants to give feedback about their interaction with our system. The questions included the intended purpose of the application and its users, contextual establishment of the virtual scene, the overall immersion, the possibilities for this tool in film, and the value of speed over precision.

Participants were given this modified VRUSE questionnaire, along with the open ended questions, immediately following the ethnographic interview.

4 Analysis

Data collected by our VRUSE questionnaire received a mean score for each question, which was calculated using Qualtrics. Then the data was exported into a spreadsheet for analysis. For each usability and diagnostic factor, mean scores and standard deviations were calculated for each section separately. We modeled our calculation analysis on a publication from IEEE 2008 also using a modified version of the VRUSE. [1]

Table I
Mean and Standard Deviation for Overall Usability Factors Scores

Usability Factors	Max	Min	Score ± SD
Functionality	5	1	3.92 ± 0.87
User Input	5	1	3.76 ± 0.88
System Output	5	1	3.73 ± 0.81
Consistency	5	1	3.88 ± 0.91
Flexibility	5	1	3.45 ± 0.88
Reliability	5	1	3.39 ± 1.03
Immersion	5	1	3.76 ± 0.86
Overall System	5	1	4.07 ± 0.68

Table II
Mean and Standard Deviation for Overall Diagnostic Factors Scores

Diagnostic Factors	Max	Min	Score ± SD
Ease of Use	5	1	3.67 ± 1.01
Appropriateness	5	1	3.98 ± 0.77
Learnability	5	1	3.93 ± 0.93
Functionality	5	1	3.86 ± 0.82
System	5	1	3.36 ± 0.69
Intuitiveness	5	1	3.90 ± 0.74
Presence	5	1	3.64 ± 0.92
Quality	5	1	4.00 ± 0.58
Input Sensitivity	5	1	4.14 ± 0.38

5 Discussion

The limitations of our research include the hardware limitations of Google's Project Tango. The ideal range of use for this device is from 0.5-4m from the device to any surface. [4] Since this device is still in development and is a prototype version, there are a few limitations that include a short battery life, overheating of the device, and crashing of applications during extended periods of use. It is also important to consider the brightness of the environment due to the infrared sensors being washed out by bright lights. Additionally, since Project Tango is Android-based, our application is limited to the Android development environment.

We can observe the response of our participants from the VRUSE Usability Survey. Because of the question correlation of the VRUSE between usability factors and diagnostic factors, we can imply user feedback based on the results. Generally, our user feedback from this survey was consistent. There were no usability or diagnostic factors that dipped below the neutral threshold, into the range of disagreeability.

5.1 Usability Factors

For VRUSE usability factors, we note that the most agreeable response of our users came from the Overall System section. This section was the only one with a combined mean greater than 4, with a score of 4.07, indicating that the subjects had a pleasurable use of the system as a whole compared to any individual sub-set factors. The usability factor with the lowest score was reliability with a mean score of 3.39. We attribute this lower score with the frequent crashing and overheating of the device due to it still being in product development. When the system had to be reset for our users to continue to interact with the device, their perception of reliability of the prototype was reduced.

5.2 Diagnostic Factors

If we focus on the VRUSE Diagnostic Factors Analysis, we found success in our project based on the high mean score in the functionality and learnability fields. Our participants averaged a score of 3.86 and 3.93 respectively, which is relatively high compared to the other diagnostic factors. This indicates our participants are aware of the purpose and main function of the device. Both Quality and Input Sensitivity means scored above the agreeable level of 4; however, we must note that these two diagnostic factors are present for only a single question in our customized VRUSE survey, meaning concrete conclusions cannot be drawn about these two specific factors. We have found that the diagnostic factor of system performance was notably lower than the other factors with a mean score of 3.36. Similar to the usability factors section, we are attributing this lower score with the instability of the device itself. In our interview process, we determined that some other reasons for this score can be attributed to image distortion of the actor in the virtual scene that can make the actor look ungrounded in the set.

5.3 Continued Discussion

We are able to determine the success of our project from the open-ended questions asked at the end of our modified VRUSE questionnaire. Our participant's dialogue affirmed the desire for pre-visualization tools for film production. In most cases, participants that had primarily avoided working with green screens and visual effects claimed they felt inspired to work with virtual assets after being introduced to our pre-visualization system. Participants explicitly claimed that this tool might be valuable to the industry as well as their own work by giving confidence to directors and actors when blocking shots in the virtual scene. We determined that immersion can be enhanced by fixing distortion of the image on the display of the device which is due to the fact that the virtual camera points straight forward and the device camera has a slight incline, causing the actor in the scene to appear ungrounded. These evaluations gave us the necessary insight into factors that we were unable to expect from the start but we can now account for in order to make this tool completely accurate and useful to any film production.

6 Conclusion

We have developed a low-cost system that allows indie filmmakers to pre-visualize their virtual assets in the on-set production process in order to save time and cost in their virtual productions. With this pre-visualization system, we will assume application will enable indie filmmakers to work with virtual assets, broaden their possibilities during on-set production, and save the production team time in the post-production process.

For further testing, we plan to test industry filmmakers rather than film students due to their increased involvement with fast paced production sets. Rather than testing in our university-provided green screen rooms, we will better represent indie film set environments by visiting existing indie film studios and conducting contextual inquiries and ethnographic interviews in their familiar space within the context of an on-going film production.

We intend to continue this project for making the pre-visualization process a valuable and rapidly-conducted step towards the production of films. We will continue to establish mobility, flexibility, and efficiency of our application. We also plan to expand the scope of the current state of the tool for full functionality of pre-visualization such as scene creation, customization, environment uploading, and fully functional virtual lighting and color-correction. We believe every film member should be contextually aware and be able to explore the possibilities of the film with a relatively simple mobile device.

Acknowledgements

We want to thank George Takahashi from Purdue University, who provided insight and expertise that greatly assisted in the product development and the research. Special thanks to Kevin Thang for major contributions as well as Esteban García, Dan Anderson, and Noah Bannister for all their help throughout this process.

References

[1] Fitzgerald, D., Trakarnratanakul, N., Dunne, L., Smyth, B., & Caulfield, B. (2008, August). Development and user evaluation of a virtual rehabilitation system for wobble board balance training. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE* (pp. 4194-4198). IEEE.

[2] G Briand. (2014, November). On-Set Previsualization for VFX Film Production. *International Broadcasting Convention*. 1-8.

[3] Kalawsky, Roy S. (1999, February). VRUSE - a computerised diagnostic tool for usability evaluation of virtual/synthetic environment systems. *Applied Ergonomics, 30*, 11-25.

[4] Milgram, P. (1994). A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information Systems*, E77-D(12), 1-15.

[5] Nitsche, M. Experiments in the Use of Game Technology for Pre-Visualization. In *Future Play '08*, pages 160-165. ACM SIGGRAPH 2008.

[6] Project Tango API Overview. (2015). Retrieved November 22, 2015, from https://developers.google.com/project-tango/apis/overview?hl=en