Augmented Virtuality in Real Time for Pre-visualization in Film

Alex Stamm*
Purdue University

Patrick Teall†
Purdue University

Guillermo Blanco Benedicto††
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.

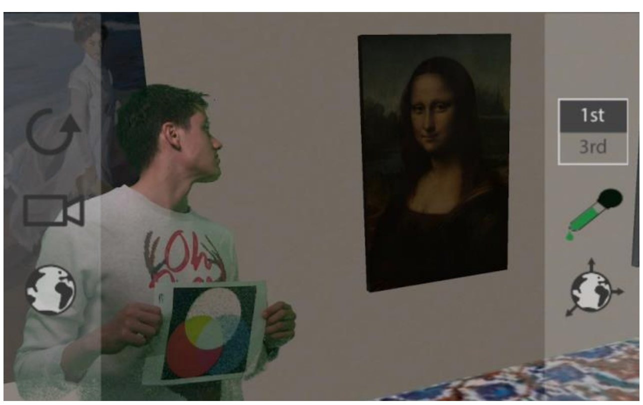

Figure 2: Actor with pre-visualization in a virtual museum.

* stamm@purdue.edu
† p.teall25@gmail.com
†† 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 based on storyboards and instruction from the director. Because this filming process must first take place in a green screen room, the problem lies in that the director must imagine this scene in his or her head. They are unable to communicate this visual to the members of the crew as well as the actors easily.

The difference between traditional film methods and using a pre-visualization tool would be that the director could create a clearer picture of what the scene might look like. Pre-visualization tools could enhance the production timeline by allowing more people the ability to see the same visualization. This would decrease the potential risks that would come if it were to look different than imagined. 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 that is 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 low-budget filmmakers who would like to use special effects in their productions. To our knowledge, there is no alternative for these low-budget filmmakers to use pre-visualization in their films due to the costly nature of these current systems. For the purposes of our research, we chose to consider independent filmmakers as any filmmaker, team or studio with a budget of less than $1M for the sake of distinction due to the lack of a consensus for how the film industry defines independent filmmakers [7]. This lack of consensus is due to the discrepancy of funding, genre and connection between the director and the studio.

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 the planning of complex motion, 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.

Our system performs the task of augmented virtuality to give power to the filmmaker to achieve pre-visualization. This system augments real-time video of real people into the virtual world of the scene of the film all from an Android tablet. See Figure 1 & 2. This is how our system falls within the realm of augmented virtuality.

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. 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].

Any mixture of real and virtual objects can be described as a location 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.

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 measurements 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

3.1 Implementation

Our system's software infrastructure is primarily built using the Unity 3D game engine with plugins provided by Google's Project Tango, our primary Android based testing device. We use the Unity game engine to put together the virtual environment and the user interface.

First, we build the virtual environment in Autodesk Maya where we model the virtual environment. Objects can be scaled later in Unity if desirable. Textures on the models are applied in Maya, acquiring UV coordinates so the texture can be mapped onto the 3D geometry. The environment is exported as an .obj file which can be imported into Unity.

The camera that we use for the virtual environment is a camera plugin from Project Tango called the Tango Delta Camera. This camera contains C# scripts written by Google that gives the Tango device the ability to track its position and perspective in real space and correlate that pose information to the virtual camera. Any change in position or angle of perspective using the Tango device in real space is automatically interpreted by the virtual camera in real time. The Tango Delta Camera is placed at origin in the virtual space and is the anchor for our frame of reference when the application first starts, and can be reset through our user interface.

The user interface is built using the UnityUI system library. The user interface as a whole was mocked up in Adobe Photoshop for the initial conceptual design. Sprite icons for the buttons were sketched and then designed in Adobe Illustrator using a square tile size to maintain consistency. The sprites were exported from Adobe Illustrator as .png's. Using the Unity UI built-in canvas and button elements, buttons were placed into positions based on the conceptual design. The sprites are then imported into Unity and are attached to the button interface.

The programming for button interaction is written in C#. Our UI script is attached to an empty game object in the Unity hierarchy of objects and in itself contains methods depending on the desired functionality. The Unity UI system library allows for C# scripts to be called upon when a user presses the button game object as an "On-Click" interaction. The built-in Unity button element is able to point to specific functions within the script if they are designated as public functions. For example, if the "Camera On" button is pressed, and the function is public, our UI script is referenced by that button interaction, and the specific function written for turning on the device's camera is called from that single button interaction.

The shaders in our application are written in GLSL, a variation of OpenGL. When the device's camera is turned on, the video seen by the user is a texture mapped onto a 3D plane quad object. Because it is a texture, it can then be affected by shader calculations. Through our chroma-key functionality, we give the user the ability to choose an input color as well as the power to alter sensitivity and smoothness as arguments. Our shader first converts the RGB values of the camera texture and the input color into a YUV color space. Then sensitivity, smoothness, and these YUV color inputs are used as arguments in a smoothstep calculation. The product of this calculation with the image's alpha, and the image's color, create transparency within the desired color range based on the central input color value. The colors outside this range remain unaffected, and these unaffected pixels are visible to the user.

Finally, our entire system in Unity is built to Android and exported onto the Project Tango development device for testing. Build settings are standard and export in the .apk format to be understood the Android device.

3.2 3D Interaction

We consulted an industry independent film professional about his understanding of virtual assets, green screens, and 3D environments so that our team could better understand the perspective of the traditional filmmaker that is unfamiliar with visual effects production. Through this process, we identified potential usability issues, the needs of the users, and how a filmmaker might interact with a user interface in a 3D space.

Through the user interface, we gave users the ability to perform 6 basic functions with our device. These functions included on/off toggles for visibility of the device camera and for the virtual world. Users also have the ability to reset the origin point of the virtual camera, change the camera perspective between 3rd person and 1st person. Lastly, our application requires the ability to manipulate the 3D virtual space and perform pixel subtraction through the process of chroma-keying. Our 3D world manipulation provides directional buttons that incrementally translates and rotates the virtual environment. Chroma-keying was accomplished by selecting a color from the visible background and having adjustment sliders for smoothness and sensitivity.

An important factor that we considered while building our system was how the user interacts with their environment. The user has the ability to manipulate the virtual environment two different ways. They can simultaneously walk through the real world, giving a correspondence to the virtual camera, and they can manipulate the virtual environment through our interface. We gave the users the limitation of single-click incremental movements through the virtual world to reduce the total amount of movement available to the user as they are already moving their perspective of the virtual set by walking through the real space.

The user interface we created exists within the virtual world space. The interface is mapped to the camera position so that the user does not notice movement in the interface panels. This gives the user a perceived 2D, static perspective of the user interface in his or her 3D virtual environment. As the user moves throughout the 3D scene, the user interface panel remains visibly constant.

3.3 Usability

During our usability research, we restricted the user to a single pre-built 3D virtual environment. We tested our application with a group of students from the film and video studies department at Purdue University and the telecommunications department at Ball State University. Our sample consisted of students, faculty, and alumni from these departments in order to represent various skill levels of student and independent filmmakers. Our participants consisted of males and females ranging from 20 to 30 years of age. The participants are familiar with mobile applications, are involved with film studies, and have experience participating in the film production process. Participants explained that they were unfamiliar with using virtual assets in their film productions.

The usability testing of our 3D interface and visualization system was conducted through a task-based contextual inquiry which provided insight into the learnability of our application, the effectiveness of the tool on a film set, and the immersion into the virtual scene. These contextual inquiries began with preliminary semi-structured interview questions about participants' film experience and their familiarity with virtual assets within our green room testing environment. Participants were then introduced to the Tango device, as well as our visualization application, and were instructed to visualize the scene and to give proper direction to the provided actor. The virtual scene provided was a single room museum with famous paintings hanging on the walls. Participants were told to align the actor within the virtual environment such that they would be observing the Mona Lisa and Persistence of Memory paintings as realistically as possible.

Due to the precision that most participants sought, time on task was not among the evaluated metrics. The metrics evaluated included task success, errors, efficiency, and learnability. Aside from our participant and actor, we had one observational note-taker and a conductor to provide instruction, all of which were fulfilled by the authors of this paper. The Tango screen, as well as the entire room, was recorded to give insight into the overall movement and direction around the set. Succeeding the completion of these tasks was an additional semi-structured interview for post evaluation. Lastly, participants were asked to complete our modified VRUSE Usability questionnaire that includes open-ended questions about the value of this technology. Upon completion of our testing, participants were rewarded with king-sized candy bars to thank them for their time and help.

3.4 Post Evaluation Questionnaire

Our post evaluation testing method was a modified version of the VRUSE usability questionnaire hosted by Qualtrics [8]. This questionnaire is a computerized diagnostic tool for the usability evaluation of virtual/synthetic environment systems [3]. The full version consists of 100 questions that combine a list of usability and diagnostic factors for determining the total usability of a synthetic environment. It is broken into ten usability factor sections that 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 such as removing the user guidance and help section which is tailored to the user being able to request help for online assistance.

VRUSE consists of 50 negative and 50 positively toned questions so as to receive a balanced response from the participant [3]. When our team created the customized VRUSE questionnaire, the reduced version consisted of 40 positively-toned questions and 29 negatively-toned questions. For each negatively-toned question, we reciprocated the value to add to the total mean. The Likert scale values remained the same for participants, but for the purposes of analysis, the scores of questions with a negative connotation were reversed to add to the total mean.

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.

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

5.1 Limitations

The limitations of our research included 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 [6]. Since this device is still in development and is a prototype version, there were a few limitations that included a short battery life, overheating of the device, and crashing of applications during extended periods of use. It was 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 was limited to the Android development environment.

5.2 Results

We can observe usability and diagnostic factors of the customized VRUSE based on the quantitative results. There were no usability or diagnostic factors that dipped below the neutral threshold, into the range of disagreeability.

For usability factors, we noted 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 participants 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 attributed 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.

We had 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, we attributed this lower score to the instability of the device itself. In our interview process, we determined that some other reasons for this score could be attributed to image distortion of the actor in the virtual scene that made the actor look ungrounded in the set.

5.3 Continued Discussion

One important position that we did not test with our research was how this system benefited the actor. One must consider the importance of this information to the actor, the person whom must "interact" with these virtual assets. While pre-visualization is important to the director and cinematographer in a production for the purpose of generating camera angles, it is likely important to the actor to know exactly where they are placed within a virtual space. This would give more believability to the actor's performance, increasing his or her contextual awareness within the scene. One could consider feeding the display of the Android device to an external monitor so the actor could see where they are located within the virtual set; however, we did not consider testing the actors as they are not the primary intended user of our system.

6 CONCLUSION

We have developed a low-cost system that allows independent 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 this application will enable independent filmmakers to work with virtual assets, broaden their possibilities during on-set production, and save the production team time in the post-production process.

ACKNOWLEDGMENTS

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

[7] V. Renée Sundance Numbers in Independent Film. (2014). Retrieved November 22, 2015, from http://nofilmschool.com/2014/01/sundance-infographic-numbers-in-independent-film

[8] The data analysis for this paper was generated using Qualtrics software, Version 2015 of Qualtrics. Copyright © 2015 Qualtrics. Qualtrics and all other Qualtrics product or service names are registered trademarks or trademarks of Qualtrics, Provo, UT, USA. http://www.qualtrics.com