Assessing the Value of 3D Software Experience with Camera Layout in Virtual Reality

Daniel Hawes
School of Information Technology
Carleton University
Ottawa, Ontario Canada
dan.hawes@carleton.ca

Rob Teather
School of Information Technology
Carleton University
Ottawa, Ontario Canada
rob.teather@cunet.carleton.ca

Ali Arya
School of Information Technology
Carleton University
Ottawa, Ontario Canada
AliArya@cunet.carleton.ca

Max Krichenbauer
Hankyu Five Anex Building 2F
1-12 Kakuta-cho, Kita-ku,
Osaka, 530-0017 Japan
max@marui-plugin.com

***Abstract*— Preproduction is a critical step in creating 3D animated content for film and TV. The current process is slow, costly, and creatively challenging forcing the layout director (LD) to interpret and create 3D worlds and camera directions from 2D drawings. Virtual reality (VR) offers the potential to make the process faster, cheaper, and more accessible. We conducted a user study evaluating the effectiveness of VR as a preproduction tool, specifically focusing on prior 3D modeling experience as an independent variable. We assessed the performance of experienced 3D software participants to those with no experience. Participants were tasked with laying out a camera shot for an animated scene. Our results revealed that the experienced 3D software participants did not significantly outperform their non-experienced counterparts. Overall, our study suggests that VR may provide an effective platform for animation pre-production, "leveling the playing field" for users with limited 3D software experience and broadening the talent pool of potential LDs.**

Keywords*— *Virtual Reality, camera, 3DVE, layout, DOP, director, layout director, cinematographer, composition, Maya, HTC VIVE, VIVE PRO, Marui, VR-Plugin

I. Introduction

VR and 3D visualization technologies are rapidly evolving and offering many advantages in design and consumer experience previsualization [10]. VR technology solutions offer the potential to improve the animation production pipeline, in turn saving time and costs for animation producers. A typical example of the studio pre-production workflow is seen in Figure 1. Within this workflow, the approved scripts and visual images are fed into a previsualization process to enable the animatic and layout processes. The scripts and visual imagery are then crafted into 2D storyboards with voices, audio effects and proposed camera moves to become "animatics". These animatics are timed out and become the creative input to the ensuing layout and animation process.

Once the LD is satisfied with the initial asset placements, she positions the camera to represent the preferred audience perspective. The process presented in Figure 1 however, has many pain points. It is time-consuming, expensive and limited in quality by the availability of talented LDs. The current process can take 3-6 months for one 22-minute TV episode and over a year for feature-length CG movies. Further, the LD must translate the 2D animatics into the 3D set, often guessing how to

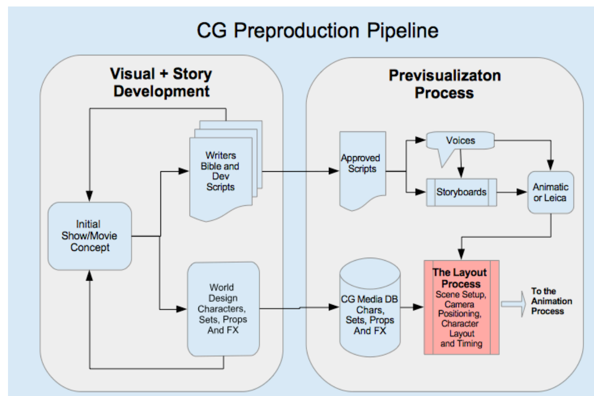

Figure 1 - Animation Preproduction Workflow used by March Entertainment in the Production of "Mia and Me" TV Series 2012

best represent a scene in 3D. Errors at this stage will propagate throughout the production. Moreover, LDs are scarce as they require both photographic acumen and specialized software expertise.

To address these challenges, we proposed to use VR as a tool for camera layout reducing time and cost from the existing process and also to determine if VR could help users with limited specialized 3D software experience perform this critical function. In doing so, we could streamline the process and reduce or eliminate a critical skill requirement and potentially draw on a larger pool of LD talent. We conducted an experiment comparing the performance of experienced Maya 3D software participants (ME) in comparison to participants with no Maya, or equivalent specialized 3D software experience (NME).

II. Related Work

The idea of employing immersive VR as a film-making tool has been cultivating in academia for over fifty years since Sutherland's "Ultimate Display" [1]. Even earlier, Knowlton considered computer techniques for animated movies with his BEFLIX movie language [2]. Subsequent research has identified innovations in storyboarding, layout, animation and camera capability with perhaps the most profound impact centering around the effects of immersion and presence.

Pausch et al. [3] studied the potential effects of immersion and presence in a VR environment compared to the same tasks performed in a 3DVE desktop environment. The tasks required locating objects within a camouflaged environment. While there was no difference in finding objects, the VR group was able to deduce the *absence* of options 41% faster than the desktop group. The researchers concluded that the enhanced mental frame of reference resulted in less redundancy in their search techniques arguably attributed to a better sense of being there, i.e., presence. Similarly, on a CG film shoot, improved presence and better mental frame of reference could enable more confidence in camera placement and movement and in less time.

Bowman and McMahan [4] highlight the potential benefits of immersion, suggesting that the most intuitive benefit is improved spatial understanding. As the human brain is highly optimized for reconstructing 3D scenes from these images, it would seem to validate Pausch's findings above [3].

In a 2018 German study, a VR animation workflow was considered as a means of improving the animation process [5]. The authors, employing an HTC VIVE system created a form of animation puppetry to animate characters. Based on expert input, they added a virtual animatable camera that could also be moved around as an object in the scene. The system was evaluated by four experts, who concluded that while the animation process worked, and the workflow was dramatically faster, precision was lower compared to traditional animation tools. Considering 3D animation tools like Maya have been evolving for over 30 years, this is unsurprising; nevertheless, it does suggest some potential for the preparatory and scene setup tasks requiring less precision but more previsualization capability than the traditional storyboarding and animatic process currently offers.

Henrikson et al. [6] argue that storyboards are quick sketches with little context, handed out on set to provide a reference for actors, directors, stunt artists, and other cast members. While this is true for live-action film, in animation workflows, storyboards are professionally drawn and animatics contain perspective, audio, and framed presentations with camera movements for every shot. A layout director could do this work in VR in a fraction of the time, potentially rendering much of the 2D storyboard process obsolete. This process could offer similar advantages with mixed reality scenarios. A 2008 study of mixed reality camera previz, successfully demonstrated an action rehearsal system that combined live action and animated 3D elements. The researchers concluded that the process could be applied to a mixed reality and even theatrical productions [7]

While the enhanced 3D visualization offers great disruptive potential, pushing the technology too far could also be counter-productive. Henrikson et al. [6] also highlighted challenges; particularly those of attempting to construct VR stories (VRS) within VR. Notably, the complexity and cost of providing multiple perspectives can be problematic; the authors suggest instead limiting the number of perspectives. Cinematic VR, for example, (CVR) [8] was proposed as an alternative to providing unlimited perspectives for VR films.

Similar concerns were first raised when developing Aladdin VR for Disney's Epcot Centre [9]. They employed bi-ocular (rather than stereo) CRT HMDs of the day. The fidelity was sufficient, as post-ride questionnaires and interviews revealed that riders of all ages suspended disbelief and found the experience to be very compelling. The authors considered possibilities like allowing guest-controlled cameras, noting that while the idea that the guest can "be the director" is compelling, it creates several challenges. First, all content (sets, characters, props) must be of acceptable viewing quality from all angles, adding considerable production value and cost. Second, the freedom could come at the expense of a more effective narrative perspective authored by a talented director. In a professional environment, however, this trade-off of control vs. constructed narrative might benefit layout directors. So while a consumer tool may not be optimal, it could provide necessary previsualization capability to professionals.

Ultimately, the fit between the tool, the medium, and the desired intent must be considered and balanced in the effort by the filmmaker to create a compelling emotional experience within VR and it all starts with the camera layout.

III. Methodology

We conducted a user study using the Marui VR plugin for Maya. We were primarily interested in determining if participants' prior experience with 3D modeling tools (Maya especially) influenced their ability to use the VR tool.

A Participants

Sixteen paid participants (9 male, 7 female) took part in the study. They were divided into two distinct groups: Maya experienced (ME, N=8, M=6, F=2) and no Maya experience (NME, N=8, M=3, F=5). Maya experienced participants had at least 1 year of sustained experience in Maya or similar 3D software. Participants with less than one-day of Maya experience were placed in the non-Maya experienced group.

B Apparatus

We conducted the experiment using a Windows 10 workstation with the following specifications: Intel 8700K 6-core, Nvidia GTX 1080 TI video card, 48 GB of DDR4 RAM. The VR system was an HTC VIVE Pro with dual AMOLED, 3.5-inch diagonal screen, full resolution of 2880x1600 or 1440x1600 pixels per eye and refresh rate of 90 Hz and a 110° field of view (FOV) with 2 VIVE hand controllers and 2 Base Stations (V1)

The 3D software platform was based on Autodesk's Maya 2018. The VR software component was provided with a Maya VR Plugin, called Marui v. 3.0.5. The plug-in runs within Maya by creating an additional Maya viewport (3D view and editor window), which is visible in the VR headset. Thus, it stays consistent with the normal Maya user interface and scene display as in Figure 2. By using the OpenVR SDK, MARUI calculates the position and view parameters of the HMD and manipulation of the VR controllers inside the Maya scene and translates VR user interactions into Maya operations which are immediately applied to the scene. Optionally, additional Maya windows are displayed in the VR environment as floating windows. Thus, it is possible to see the 2D rendering of Maya cameras in VR while manipulating the cameras in real-time.

C Procedure

The experiment started by presenting two short videos to each participant. The first video demonstrated how the standard camera layout process worked in Maya. The second video presented the new VR process employing Marui. We gave participants very specific instructions to indicate *how* to use the controls with Marui. After completing the videos and signing the consent forms, we provided a 15-minute hands-on training session with the HMD and controllers to reinforce the proposed camera layout task as viewed in the VR video.

Figure 2 – Participant performing layout, viewed in Maya Viewport

Participants then proceeded to the testing area where they were presented with a 12-second, pre-animated 3D scene created in Maya. Two animated characters were modeled, rigged and posed in a living room talking to each other. There were no previous camera setups in the scene, simply the standard Maya viewport perspective as seen Figure 2.

Participants were then asked to enter the VR environment by donning the HMD and to ensure the controllers were visible and properly aligned. We ensured that the HMD was comfortable and that they were fully immersed in the scene. From the Maya viewport, we were able to see exactly what the participant was viewing in the HMD. Once we confirmed that all was OK and that they understood the task, we started a stopwatch and began recording the session. The participants then started the process of creating, configuring and setting up the camera, adding a viewport, and manipulating the camera to achieve an interesting perspective on the character or characters. We stopped testing at the 5-minute mark as we felt that requiring more than five minutes would be an indication of confusion or other problems; most participants completed well before this.

Two judges, both experienced directors with multiple professional TV/film animation credits, viewed the VR performance measuring time and the final aesthetics of the framed shot. At the end of each session, the judges briefly conferred and awarded a proficiency score between one and ten based on an equally weighted combination of the time and quality. Upon completion of the test, participants were asked to complete a UX questionnaire and some follow up demographic questions. They received a $50 cash payment upon completion.

There were several steps required to setup, configure, select, manipulate and frame the opening shot. An example of how to perform each of these steps was presented to the participants in the instructional video prior to the testing. These steps, ranged from initial alignment to the final camera layout.

D Design

Our study employed a between-subjects design with a single independent variable, Maya experience, with 2 levels: Maya-experienced (ME), or no Maya experience (NME). The main performance-dependent variables included completion time (seconds), and quality creating an overall proficiency score. Proficiency scores were ranked between 1 and 10 based on the judges score, with a half-point penalty for poor speed, and a half-point bonus for good process. We also surveyed several UX-related dependent variables including perceived effectiveness, ease of use, nausea, presence and several other related variables rated on 7-point Likert scale.

IV. Results

As expected, the ME and NME groups performed comparably in terms of task completion time, see Figure 3. An independent-samples t-test confirmed that the difference in completion time between the ME and NME groups as not significant ($t_{14} = 0.284$, $p = .78$).

We also compared the judge-determined proficiency scores between the two groups. According to an independent samples t-test, proficiency scores were also not significantly different between the groups ($t_{14} = -0.913$, $p = 0.37$). The mean proficiency scores are seen in Figure 4.

The NME group did have slightly higher scores but since the result was not significant, we cannot take this as definitive evidence that the NME group did better, but it is encouraging nonetheless.

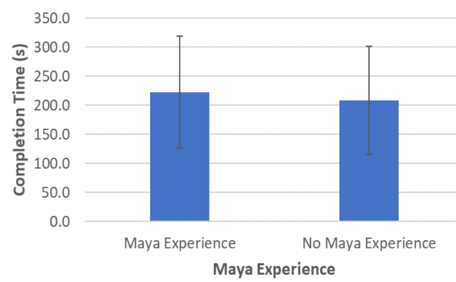

Figure 3 - Mean completion Time. Error bars show ±1SD

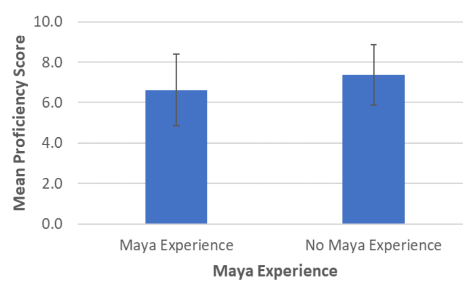

Figure 4 - Mean proficiency scores for both groups. Error bars show ±1SD

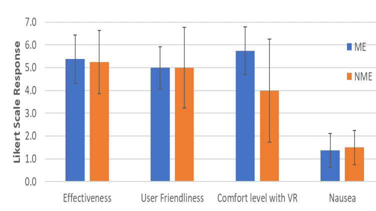

Figure 5 - Mean proficiency scores for both groups. Error bars show ±1SD

On a 7-point Likert scale questionnaire, participant responses were generally high, see Figure 5. Both groups scored above 5 on both perceived effectiveness and user-friendliness of the VR solution. While comfort levels in VR for the NME group were at the midpoint, the overall average was almost 5 (4.9) suggesting that participants overall perceived the VR solution as both effective, user friendly and within a reasonable comfort zone. The nausea scores were very low suggesting that it was not a problem.

Feedback from both groups was positive with constructive feedback focused on issues related to the complexity of the controllers and the need for more training time. For example, participants noted "There was some initial confusion but it became easier with practice. The UI could use enhancements". In general, participants saw great potential for the solution recognizing the limitations in such a short test. Another noted, for instance, "The motions are harder at first, but easy once grasped. It does jump around a bit. VR is the future". Several participants commented on the learning curve, noting the need for further practice/experience with the VR tool.

V. Discussion and Conclusion

The results of the experiment are positive and suggest that VR applications like Marui have the potential to inform process redesigns to make the preproduction workflow faster and less costly. Specifically, our results suggest that the use of VR in this context accelerates the learning curve relative to the conventional 3D modeling process. We found that novice participants with no prior 3D modeling experience were *not worse* at a camera layout task than experienced modelers. We surmise that VR may allow talented visual storytellers from other domains (e.g., live-action or gaming) with less domain-specific experience to perform these cinematic functions, potentially broadening the talent pool.

While observing the user tests, the director/judges noticed and documented a recurring set of UX challenges that participants were experiencing. First, many of the participants struggled with the camera selection function because of object occlusion (i.e., viewport) or uncertainty status when invoking the quick-drag feature. When establishing the final shots, focal lengths seemed to be changing randomly and while they can be adjusted, few participants were able to do this even though it was presented in the instructional video.

Although we recommended setting up camera size and aspect ratios before creating the camera, this option was either forgotten or missed. We assume partly because the settings were on the top right of the POV camera button, not in clear view and a bit awkward to select. Occasionally, when they were confused or could not immediately see the camera, participants would keep pressing the camera POV button, creating multiple cameras and causing further confusion.

Finally, the controller functions were a bit sensitive and some of the participants would accidentally put pressure on the side buttons, unintentionally moving the scene around and changing the user location relative to the scene. While the Maya plug-in was impressive, informal discussions centered around how the experience could be improved. In general, design recommendations included.

- Enabling multi-person collaboration, which would allow an LD, camera person, and even an asset coordinator to work together to do camera layout and set design.

- Hand control functions that could employ simpler selection techniques, camera setups and change functions.

The preproduction process is a critical aspect of CG production. It is within this process that the visual story is architected, creating a digital blueprint for production. The process is labor intensive, costly and limited by the quality and availability of specific technical and creative talent.

The results did suggest that experienced Maya users did not demonstrate any advantage over VR users with no Maya experience in performing camera layout functions. Further, our testing revealed insights suggesting and validating previous studies suggesting VR tools like Marui represent the future for animation previsualization tasks.

We also provided several key UX insights and two design recommendations to guide the evolution of Marui, or similar VR layout tools. Finally, we propose that the results did represent a small sample size and future studies in this area increase both the scope, and frequency of participants to gain more insights into how to best apply VR technology within the animation production workflow.

Acknowledgment

Thank you to all of the participants, especially the students from Conestoga College, Todd Janke, Program Director and Professor Edin Ibric of Durham College who served as Animation/Maya expert. Finally, thanks to Marui's Max Krichenbauer for the support and feedback during our testing process.

References

[1] Sutherland, Ivan E. "The ultimate display." Multimedia: From Wagner to virtual reality (1965): 506-508.

[2] Knowlton, Kenneth C. "The Beflix Movie Language." *Proceedings of the Spring Joint Computer Conference*. 1964.

[3] Pausch, Randy, Dennis Proffitt, and George Williams. "Quantifying immersion in virtual reality." (1997).

[4] Bowman, Doug A., and Ryan P. McMahan. "Virtual reality: how much immersion is enough?." Computer 40.7 (2007): 36-43.

[5] Vogel, Daniel, Paul Lubos, and Frank Steinicke. "AnimationVR-Interactive Controller-Based Animating in Virtual Reality." 2018 IEEE 1st Workshop on Animation in Virtual and Augmented Environments (ANIVAE). IEEE, 2018.

[6] Rorik Henrikson, Bruno De Araujo, Fanny Chevalier, Karan Singh, Ravin Balakrishnan. Multi-Device Storyboards for Cinematic Narratives in VR

[7] Ichikari,, Ryosuke, Ryuhei Tenmoku, Fumihisa Shibata, Toshikazu Ohshima, Hideyuki Tamura, "Mixed Reality pre-visualization for filmmaking: On-set camera-work authoring and action rehearsal." Int. J. Virtual Reality 7, no.4 (2008): 25-32.

[8] Green, Melanie C., and Timothy C. Brock. "The role of transportation in the persuasiveness of public narratives." Journal of personality and social psychology 79.5 (2000): 701.

[9] Pausch, R., Snoddy, J., Taylor, R., Watson, S. & Haseltine, E. 1996. Disney's Aladdin: first steps toward storytelling in virtual reality. In Proc. SIGGRAPH '96, 193-203.

[10] MH&L (2015, November 12) Is virtual reality coming to the supply chain? Retrieved from http://www.mhlnews.com/technology-automation/virtual-reality-coming-supply-chain