

# VR as a Content Creation Tool for Movie Previsualisation

Category: Application



Figure 1: We propose a novel VR authoring tool dedicated to sketching movies in 3D before shooting them (a phase known as *previsualisation*). Our tool covers the main stages in film-preproduction: crafting storyboards (left), creating 3D animations (center images), and preparing technical aspects of the shooting (right).

## ABSTRACT

Creatives in animation and film productions have forever been exploring the use of new means to prototype their visual sequences before realizing them, by relying on hand-drawn storyboards, physical mockups or more recently 3D modelling and animation tools. However these 3D tools are designed in mind for dedicated animators rather than creatives such as film directors or directors of photography and remain complex to control and master. In this paper we propose a VR authoring system which provides intuitive ways of crafting visual sequences, both for expert animators and expert creatives in the animation and film industry. The proposed system is designed to reflect the traditional process through (i) a storyboarding mode that enables rapid creation of annotated still images, (ii) a previsualisation mode that enables the animation of the characters, objects and cameras, and (iii) a technical mode that enables the placement and animation of complex camera rigs (such as cameras cranes) and light rigs. Our methodology strongly relies on the benefits of VR manipulations to re-think how content creation can be performed in this specific context, typically how to animate contents in space and time. As a result, the proposed system is complementary to existing tools, and provides a seamless back-and-forth process between all stages of previsualisation. We evaluated the tool with professional users to gather experts' perspectives on the specific benefits of VR in 3D content creation.

**Index Terms:** Human-centered computing—Visualization—Visualization techniques—Treemaps; Human-centered computing—Visualization—Visualization design and evaluation methods

## 1 INTRODUCTION

Film and animation industries rely on a range of techniques to prepare and prototype movies before shooting them. Hand-drawn storyboards or digital storyboards [8] remain one of the most traditional ways to design the visual and narrative dimensions of a film sequence.

With the advent of realistic real-time rendering techniques, the film and animation industries have been extending storyboards by exploring the use of 3D virtual environments to prototype movies, a technique termed *previsualisation* (or *previs*). Previsualisation consists in creating a rough 3D mockup of the scene, laying out the elements, staging the characters, placing the cameras, and creating a very early edit of the sequence. As an addition to the storyboard, the previsualisation stage provides a mean to creatively explore multiple framing or editing choices, and to confirm or confront initial design ideas from early stages with layout, shooting and editing constraints. Most large budget movies nowadays rely on

previsualisation technologies as a way to reduce the costs by early anticipation of issues that may occur during the shooting.

Previsualisations of movies or animated sequences are actually crafted by dedicated companies composed of 3D artists, using traditional modelling/animation tools such as Maya, 3DS or Motion Builder. Creatives in film crews (editors, cameramen, directors of photography) however do not master these tools and hence can only intervene by interacting with the 3D artists. While a few tools such as FrameForge3D, MovieStorm or ShotPro [2, 4, 6] or research prototypes have been proposed to ease the creative process [23, 27, 28], most tools remain complex to master, outsourcing the control of the content, and hampering the creativity of film crews. In such a context, we propose to fill a gap by designing accessible and dedicated content creation tools at the storyboard, previsualisation, and technical stages so that part of the creative power can be placed back in the hands of film creatives, and contents can be co-created and co-reviewed by creatives and 3D artists.

The founding idea in our approach is to rely on VR interactive technologies to author these animated 3D contents, building on the immersion offered by VR together with the capacity to manipulate contents using 6D input devices.

In this paper we present how we addressed specific issues in previsualisation, and report feedback from expert users. Our approach extends techniques such as the Directors Lens [23] and [22] to not only manipulate cameras, but also manipulate scene layouts, 3D character poses and animations, as well as complex camera rigs (cranes) or light setups. Through experiments that involve expert users only, we demonstrate the strengths of this approach, namely (i) the capacity to be easily integrated into existing pipelines, (ii) the improved creativity of users, and (iii) the capacity to seamlessly navigate between all the different stages. To the best of our knowledge, this paper represents the first VR tool that covers all the previsualisation previsualisation stages and received very strong interest from film schools and experts. Leading animation (Pixar, BlueSky studios, Disney) and VFX companies (MPC, ILMx, Halon) have just started to explore the possibilities of VR creation in their pipelines, and discussions with these companies during FMX and Siggraph 2018 showed the clear advance our approach has.

In the following sections, we will report progress in authoring tools for the film industry and will also focus on VR/AR as content authoring tools (Section 2). We will then rely on discussions with experts to derive design considerations (Section 3), before describing our approach (Section 4). Experiments are then reported in Section 5 before drawing conclusions (Section 6).

## 2 RELATED WORK

We first explore the use of VR as a mean to create contents interactively rather than as a way to interactively experience contents. From there, we report and comment on tools dedicated to storyboarding and previsualisation before discussing the range of techniques that rely on VR and AR technologies for the design of cinematographic contents.

### 2.1 Virtual Reality as a Content Creation Tool

A number of drawing and authoring tools rely on VR technologies to create artistic contents such as TiltBrush [7] or Oculus Story Studio [5]. Tracked controllers are used as brushes to interact and paint in space. Despite their obvious benefit in creating 3D contents, such tools are not suited to the design of storyboards (as reported in [18]), nor to the design of previsualisations (due to the amount of work required to redraw characters in different poses throughout the sequence). Given the ease of use in VR manipulations, VR interfaces have been explored for general sketching tasks [15].

### 2.2 Tools for Storyboarding and Previsualisation

Multiple digital storyboarding tools are used by the film and animation industries (too many to actually provide an exhaustive list). Essentially these tools support artists in creating and editing drawn sketches. For example, Henrikson et al. [19] proposed a system to design stereo storyboarding. However, tools nowadays stretch towards more animated contents (see ToonBoom [8]) and start to integrate 3D scene layouts and camera placements. While intuitive for artists, these tools remain complex for non-specialists, and interaction controls of 3D contents are very similar to those found in most 3D modellers. Systems have also been proposed to rehearse theater plays using virtual environments [29].

In reaction to the complexity of 3D modellers and to address the specific needs of the film industry, a collection of high-level design tools have been proposed to create previsualisations ([2,4,6]). While the interfaces have been simplified and specialized, the control of 3D contents (mainly positioning and orienting them) remains a task that requires to switch between different interactors, and switch between camera views to control the spatial relations between entities (non-immersive systems).

Researchers have also been exploring the use of physical mockups through miniatures [20] – a technique used in the film industry to control complex camera rigs. In contrast, by proposing a VR immersion and tracked controllers, our system provides a straightforward control of entities (objects, characters or cameras).

### 2.3 AR and VR in storyboarding and previsualisation

Within the specific space of design tools to support the creation of cinematographic contents, a number of approaches have been explored. Such approaches can be characterized according to the stage at which they intervene in the production workflow.

**Storyboarding.** The use of VR in storyboarding has mostly been explored with the intent of creating narratives for VR. Shin et al. [28] proposed an AR system to author contents, mixing real video footage and 3D generated footage. Henrikson et al. [18] propose an approach to the creation of VR narratives by mixing traditional devices such as tablets to draw storyboards, with VR to experience and validate narratives. While in essence, the approach provides a means for a closer collaboration between drawing artists and filmmakers, the expertise of professional artists is still necessary. In opposition, our approach consists of designing tools that do not require 3D expertise.

**Scene layout.** Well known tools such as the VR Editors of Unreal and Unity3D propose a range of effective interactors to handle the creation and manipulation of 3D environments (designing, transforming and positioning entities in 3D). To this end, the editors rely on classical interaction metaphors for 3D manipulation [10].

While we build on these classical interaction metaphors to manipulate the elements composing the scene, we extend the approach by proposing character manipulation using inverse kinematics, camera manipulation and editing.

**Camera framing in virtual environments.** As a mean to regain control over the creative process, a number of filmmakers have pushed the idea of Virtual Camera Systems, *i.e.* devices tracked in position and orientation that enable the control of a virtual camera in a virtual 3D environment. Such systems provide the filmmakers with means to creatively explore camera angles, scene layouts and character layouts [1]. The use of virtual camera systems has been later extended with the possibility to navigate between a collection of automatically computed views on the scene and create an edit of a filmic sequence. Our approach relies on these ideas, but extend them by using VR immersion and tracked controllers to address a wider range of stages (creating a storyboard, laying out the scene, and controlling character poses).

Researchers have also been exploring questions of editing in VR [27] for 360 videos. Mixed reality technologies were exploited in the specific case of previsualisation. Ichikari et al. [21] designed an AR system to author camera placement in a mixed reality context, by adding virtual contents to real footage. The system relied on a fixed camera with rotational encoders (pan and tilt) which fed rotational information to synchronize the real and virtual camera motions. The system, however, remained limited to the specific task of exploring camera framing and staging virtual contents in a real background.

Later, technologies referred to as on-set previsualisation, enabled the real-time rendering of virtual contents in synchronization with real contents to anticipate the effect of visual effects in scene rehearsal [11]. Typically, by coupling camera tracking techniques with chroma keys, green screens can seamlessly be replaced by virtual contents. Camera tracking technologies rely on visual odometry or single localization and mapping techniques [25], or visual tracking of static targets [12]. These approaches are more dedicated to explore and validate camera angles rather than interact with the scene's virtual contents. Recently the previsualisation tool ShotPro has proposed an augmented reality mode that enables the creation of virtual scene layouts within a real environment, while also using visual odometry to creation camera angles and shots.

Closer to our work, [26] and [13] started to run exploratory experiments to evaluate how both novices and non-technical professionals from the film, animation, and theater domain assess the use of VR for previs. Early results show very positive feedback that encouraged us to push this research topic further in this paper.

## 3 DOMAIN ANALYSIS

Digital technologies have largely impacted the process of creating real or animated movies. Ubiquitous in the later production stages (*eg.* for visual effects and compositing), digital technologies are now gaining importance in the pre-production stages, *i.e.* in preparing the production. Typically, previsualisation is evolving into what will eventually become a mandatory stage in the authoring of movies. Currently required by producers for preparing high budget movies, previsualisation is more and more used to rehearse complex scenes in TV shows and medium budget productions.

In order to have an in-depth understanding of the authoring process in storyboarding and previsualisation, and what would be the requirements for novel tools, we organized informal discussions with two film practitioners who are also teachers in a leading film and animation school, two technical directors and two artists from previsualisation companies. The process is the following. The director first analyses the movie script<sup>1</sup> to create a *technical breakdown*

<sup>1</sup>A script is a textual document that describes the scenes without providing technical information as to which camera angles, scene layouts or cut between camera angles to perform.

of a sequence. The technical breakdown identifies which type of scene, camera and character layouts will best translate the script intentions. The technical breakdown is then converted into a *storyboard*, a process by which artists in relation with the film director create a visual representation of the script through drawings or digital tools [8]. The director iterates on the storyboard with artists, and on completion, a detailed version of the storyboard (called clean-up) serves to communicate the directors' intentions to the film crew.

The *previsualisation* stage then consists in working from the storyboard representations to create a 3D mock-up version of the scene with decors (background), layouts (how the different characters are placed in relation with the scene) and cinematography (deciding on camera angles and camera motions). Previsualisation also helps to create a first *edit* of the sequence, determining its rhythm and the continuity between the cameras. Finally this previsualisation stage (that is essentially focused on the narrative and aesthetic design of the sequence) is generally followed by a technical stage (called techvis), in which technical information is added, to communicate with the film crew (camera lenses, spatial organisation of the scene, types of camera rigs and light sources). Nowadays, previsualisation is performed through general purpose 3D modelling tools (Maya, MotionBuilder, 3DS) or dedicated tools such as Moviestorm [4], ShotPro [6] or iClone [3]. In all cases, 3D artists are in charge of these stages, working from the storyboard and through iterations with the film crew.

### Design considerations

A central issue reported by the film directors we interviewed is that these technologies make them lose control of the created contents at the early and important design stages, which are the most creative ones. All operations are performed by artists using elaborate sketching tools or 3D modellers most of which are not accessible to film directors. The iterative process through which the design is performed through back and forth with artists also limits the creativity [18].

From these discussions we derived a number of design considerations to drive our research:

**Intuitive:** the proposed tool should be easy to use by previs artists as well as members of the film crew (even for those who have limited knowledge of modelling/animation tools). Hence, the scenes can be accessed and modified at any stage by members of the film crew or previs artists. Clearly, a director would not create his whole movie using such a tool, but he should be able to intervene on any sequence, review the proposed solutions, explore and propose alternatives.

**Rapid:** the tool should enable the rapid creation and review of cinematic sequences, the exploration of multiple possibilities in a short time here being the key of creativity.

**Seamless:** the tool should merge the different stages that are the storyboard, the previsualisation, and the technical preparation of the shooting. Offering the possibility to seamlessly navigate between these stages that are traditionally organized in sequence, and generally involve different digital tools, would offer a mean to easily revise some early choice and explore multiple possibilities. In particular, technical constraints from the shooting stage (eg type of camera rig used) influences strongly the aesthetic choices made in the previsualisation stage. Hence encompassing these constraints early in the process presents a clear benefit.

**Workflow-compatible:** while all of the interviewed people were positive when discussing the idea of a VR tool to address these stages, they insisted in the capacity to be compatible with main tools in the industry, so that at any stage the created contents could be shared between the VR tool and Motion Builder.

## 4 DESIGNING AN INTUITIVE VR PREVISUALIZATION TOOL

We designed the tool around three main components, namely (i) the storyboarding mode, (ii) the previs mode (animating character, ob-

ject and cameras), and (iii) the techvis mode (placing and animating complex camera cranes and light setups). All the technical solutions proposed in our approach to address these aspects rely on traditional 6DOF manipulation techniques, but they integrate constraints specific to the tasks, to guide the user in their interactions. This paper does not innovate in the way interactions are performed, but in how practical issues from a specific domain can be easily addressed using VR technologies through constrained interaction, and how these can impact the users ease of use and creativity.

All three modes are displayed in Figure 2.

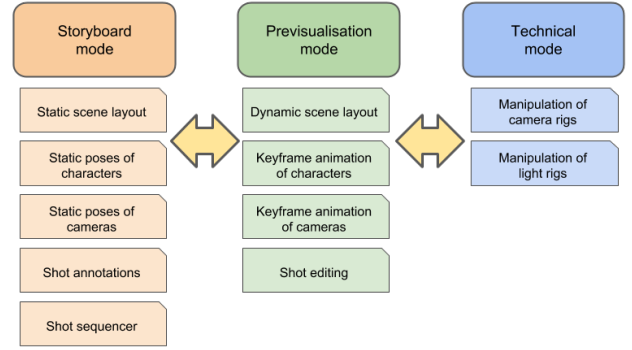


Figure 2: The three stages in the VR tool with the features developed in each stage.

### 4.1 The Storyboarding Mode

Storyboarding mode provides a mean to lay out a static scene, take static shots (thumbnail stage), and annotate the shots with elements of the storyboard visual grammar defined the intended motion of the camera (traveling, panoramic, zoom, etc. – see Figure 6) to create a content that can be shared with the film crew (clean-up stage).

#### 4.1.1 Laying out the scene

The first step consists in defining the location unit – *i.e.* indoor, outdoor, city, countryside. The availability of 3D models through asset stores provide elements that may match the storyboard, yet at this stage, storyboards generally work with grey-box environments, as a number of choices have not been fully confirmed. Different props used in the scenes can be found on the store (this exploration stage actually doesn't require a precise contents).

Users can interact with the 3D environment and our system using two spatially located VR input devices (Vive controllers in this case). We rely on classical selection, manipulation, navigation and application control techniques from the literature [10]. The right controller is dedicated to scene editing. Any object in the scene can be selected and displaced by placing the right controller inside it, pressing the trigger and dragging it in position and orientation – such classical interactors ensure quick assimilation of the system. Figure 3 shows examples of two user interactions, a raycasting metaphor to place objects newly created, and a virtual hand metaphor to enable object manipulation. The left controller is dedicated to menus and navigation. The user displacement is easily performed through a ray-casting technique, displaying a ray between the left controller and the intersection point with the scene; teleportation is performed when pressing the left trigger. The user can also fly through the environment using the controller to guide the movement; the motion starts when the user presses a dedicated button and stops when releasing the button. This last feature enforces short acceleration and deceleration period to avoid motion sickness. Finally, as illustrated in Figure 4, interactions with the menus can be performed in two ways: (i) using physical buttons on the controllers or (ii) using a



point and click metaphor with a laser directly on UI elements in the scene. Menus are used to navigate between modes (*i.e.* storyboard, previsualisation and techvis) and give access to the specific features of each mode (*eg.* add or delete an object, control time, add camera keypoints, etc.)

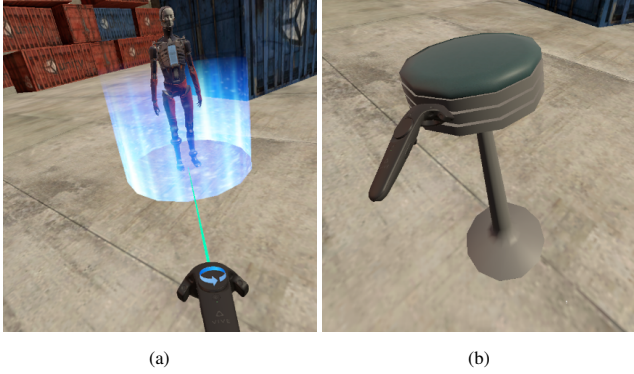


Figure 3: Users can quickly craft a scene layout by (a) adding and (b) moving objects or actors in the scene.

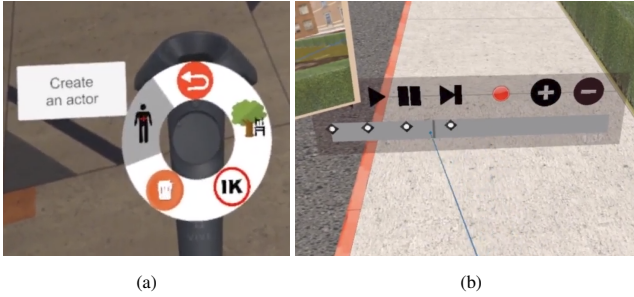


Figure 4: Users can interact with the system through (a) menus on the controller or (b) 3D UI elements in the scene.

#### 4.1.2 Controlling the pose of virtual characters

Controlling a character pose (*i.e.* manipulating joints in a skeleton) is a complex task when using classical modelling tools. Translation and rotation visual widgets are required to assist the user, and specify in which dimension the manipulation occurs and on which axis, due to depth ambiguity. Dedicated non-immersive tools [4, 6] provide a set of predefined poses (*e.g.* standing, sitting, running, etc.), but provide limited means to manipulated poses.

This paper proposes a constrained manipulation metaphor to deal with character poses in VR. No visual widgets are required as 6D manipulations are straightforward with 6D controllers and stereo vision resolves depth cues. The nodes of the skeleton can be selected as any other object in the scene. User manipulations are however designed to be constrained by the kinematic chain between the selected node and the root node of the character. Figure 5 shows some example of poses crafted with our tool.

A classical inverse kinematics solver [9] (i) constrains the maximum displacement of node given the distance and angular constraints expressed on the skeleton, and (ii) updates the positions and orientations of the nodes between the selected node and the root skeleton.

**Camera placement** Using 6DoF tracked devices on hands to handle the placement of cameras presents a tremendous advantage compared to traditional tools. While two-dimensional interfaces



Figure 5: We propose a novel interacting metaphor to craft character poses with Inverse Kinematics to ensure the naturalness of the pose. Starting from an initial pose (Left), the user can create new poses through a small set of operations.

require to mix 2DoF manipulators coupled with keyboard combinations to control 6 DoF, VR environments offer a simultaneous and natural control of all six DoFs. Once created, a virtual camera can be grabbed like any object in the scene. A preview screen and a slider to adjust the focal length are placed on top of the cameras and are always oriented towards the user as a billboard. Additional features were added from discussions with experts such (i) the possibility to detach and move the preview screen around so as to view the shot from another location, (ii) the possibility to easily switch between the previews of different cameras.

#### 4.1.3 Creating a storyboard

The storyboarder tool relies on (i) a camera manipulator to create snapshots of the scene (ii) a camera motion annotation to specify the primitive motion followed by the camera, and (iii) a monitor used to preview and organize the shots. Figure 6 shows an overview of this *Storyboarder*. A 3D model of a camera is used to control the viewpoint; it can be handled like any other object in the scene. The user can preview and modify in real-time the current camera view (including the focal length). The type of camera motion associated with a shot is specified through an overlay on the camera rendering. It follows storyboard visual conventions taken from the literature [16]. Each type of camera motion is described with a specific representation and displayed on a side panel. Figure 7 illustrates some of these representations.

Snapshots can be added, moved or deleted in the *Storyboarder* monitor. Finally, the *Storyboarder* compiles the set of snapshots together to generate a full storyboard (see Figure 8). The video provided in supplementary materials demonstrate this storyboarding mode.



Figure 6: Our storyboard interface. (Left) A camera manipulator used to define the shot, the camera focal length and specify the type of camera motion. (Right) The 3D panel used to preview and organize the created shots.

## 4.2 The Previsualisation Mode

We propose a set of features (*cf* Figure 9) to assist the user in animating the static scenes designed in the storyboarding mode.

**Camera motions** If a storyboard was created with our tool, the user can access it and reuse the cameras placed with the *Storyboarder* as starting points. To create camera motions, the user



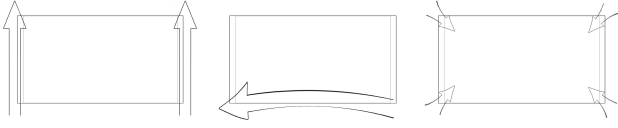


Figure 7: Our storyboard interface rely on a set conventions to describe different types of camera motions. Examples are given for (Left) a dolly up, (Center) a Pan to the left and (Right) a dolly in.

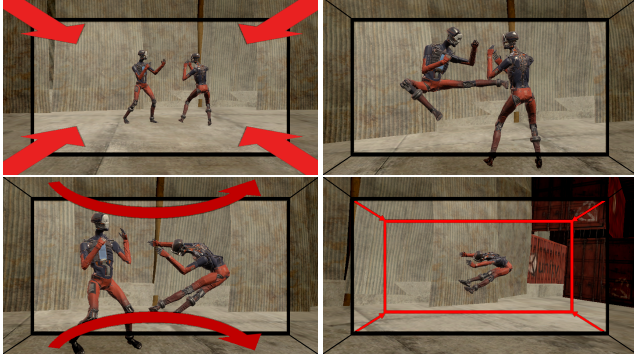


Figure 8: Examples of four different shots generated by the the *Storyboarder* in a matter of minutes.

has the possibility of (i) creating camera rails, (ii) recording camera motions by tracking the VR controllers over time, and (iii) editing the trajectories created in the first two modes. A camera rail can be designed by placing a number of control points on a spline curve representation and manipulating the positions of the control points as well as the tangents to the curve. Tangents are represented in a way that the expected interaction is straightforward leading an easy design of curve bends. The recording of tracked camera motions is a feature traditionally found in Virtual Camera Systems (see Section 2) and improved by the possibility to observe the surrounding while recording the camera motion (*eg.* to anticipate character motions). The user can insert and delete keyframes in the motion, and manipulate these keyframes.

In all modes, the camera orientation can be constrained to a given target, maintaining the position of an actor on the screen while the camera is manipulated.

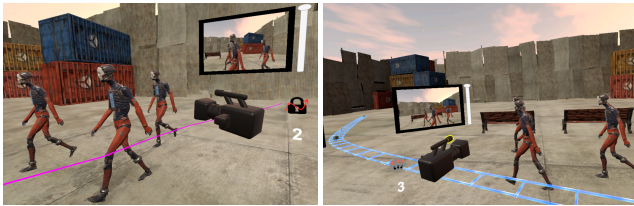


Figure 9: Example of features proposed to assist users in shooting a sequence. (Left) Trajectory recording tool. (Right) Rail crafting tool.

**Character animations** Animations are authored on characters in the same way as cameras with the possibility to (i) create key-framed animations, (ii) record animations using motion capture suits (iii) record the motions of the controller (*eg.* when manipulating an arm) and (iv) edit the motions once created whatever way they were created. In addition, the manipulations can be constrained by the kinematic chain when manipulating parts of the characters only.



Figure 10: A linear-editing interface design to create shot edits within the virtual environment

An additional panel, attached to each object being animated, displays the list of keyframes to ease their selection and manipulation.

**Editing** Finally to complete our VR previzualization tool we propose a new editing method dedicated to VR.

Standard editing tools are based on the 3-point editing principle. This method consists in inserting a portion of a camera rush, delimited by in and out timestamps (first two points), within the final edit at a specific time (third point). The method allows great flexibility such as the capacity to create flashbacks or any other type of non-linear editing. However, user interface associated with this principle is relatively complex and requires the manipulation of a wide range of precise 1DoF and 2DoF controllers mainly tailored for a precise interaction device such as a mouse. Early experiments quickly showed that importing such complex UI in the virtual world and handle interactions with a laser pointer is both counter-intuitive and counter-productive (*e.g.* limited precision in controlling a time sliders).

Thus, we proposed a simplified linear editing tool. While linear editing does not provide as much freedom as 3-point editing, it is sufficient to craft rough edits of a scene. Furthermore, the editing generated in the VR interface can be exported and later refined with traditional 2D editing tools.

The editing interface was crafted following these interaction constraints (*cf.* Figure 10). It contains a timeline to display a color-coded representation of the edited shots. The main screen displays the final edit and the two bottom screen show the initial frame of the current and next shot. As the user creates a new cut, a mosaic displays all the camera angles designed by the user. The user has the possibility to interactively change the camera angle, change the scene layout or character pose directly while viewing the results on the editing interface, seamlessly navigating between stages that are traditionally split in different tools.

### 4.3 The Technical Mode

At last the technical mode proposes means to control technical shooting devices such as complex camera rigs and light rigs, that help in preparing the real shooting of the scene.

**Control of camera rigs** A camera rig (*eg.* a camera crane) can be abstracted as a kinematic chain defined between the end effector (the camera manipulated by the user), and a root node of the camera rig, typically the position at which the camera rig is in contact with the floor. Our VR implementation considers camera rigs as any other animated entity in the scene, with the additional constraints due to the kinematic chain (limits of articular joints, limits of extensible sections of the chain). Camera rigs can be animated using keyframes.

The system supports any specified camera rig such as Technodolly Cranes – arguably the most complex camera rig in terms of kinematic chain as illustrated in Figure 11.

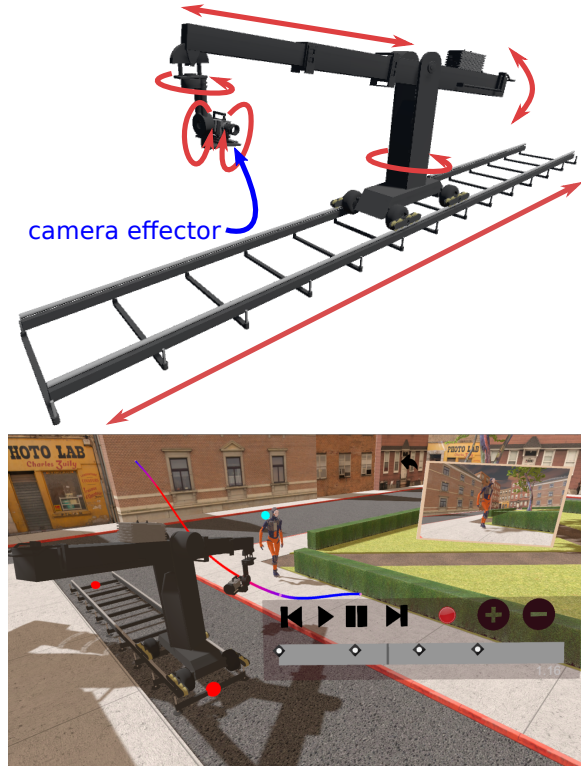


Figure 11: Our system allows the interactive and key-framed manipulation of complex camera rigs (here a Technodolly rig) defined through a list of constrained degrees of freedom (in red in the top image). The camera effector (in blue) is directly controlled by the user.

**Control of light rigs** A light rig is a structure that holds a light in a studio setup. Our VR implementation proposes the interactive manipulation of area lights through changes in position, orientation, aperture and intensity. Area lights are based on the implementation of Heitz et al. [17], and provide a rendering closer to real light staging conditions. An example is displayed in Figure 12.



Figure 12: Users can add, manipulate and control various types of light rigs available in the environment.

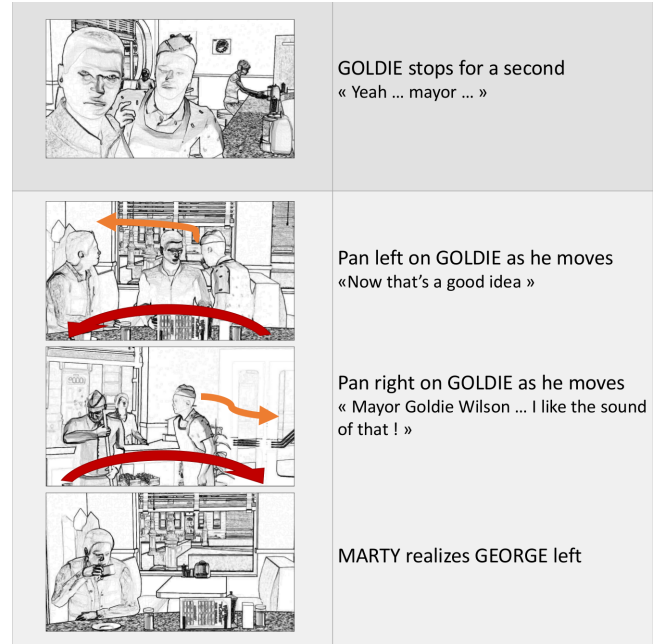


Figure 13: Images of two sequences of the input storyboard used in the first evaluation. The first sequence being a static sequence and the second involving a camera pan motion. The color arrows determined the motion of the virtual actors.

## 5 EVALUATION

This section describes two experiments, which aim at assessing the usefulness and ease of use of our tool. Our first experiment is guided by a task: users had to create an animated sequence from a given paper storyboard representation, hence reproducing realistic use case. The second experiment is not guided, so as to let users creatively design contents in all three stages (storyboarding, previsualisation and technical preparation).

### 5.1 Experiment 1: Previsualisation from a Storyboard

Thus, the first experiment aimed at evaluating how our tool can be used for reproducing a given storyboard with the goal of generating a short video (only camera and editing operations). The evaluation considers that the users do not have to modify the virtual environment nor the virtual actors.

#### 5.1.1 Apparatus and Participants

Participants were immersed in the virtual environment using an HMD (HTC Vive), interaction capabilities were provided using two HTC Vive controllers. The system was developed in Unity and ensured a constant FPS rate. Ten volunteers (all male, age  $M = 22.4$ ;  $SD = 2.9$ ), six from a local film school and four from our lab, took part in the experiment. Half of the participants considered themselves experts in cinematography, film editing, and another half novice in terms of 3D modeling and in the use of VR devices (not exclusive).

#### 5.1.2 Experimental Protocol and Design

Upon their arrival, participants were asked to sign a consent form and fill a pre-experiment questionnaire to gather relevant demographic information for the study (age, sex, previous experience with VR devices, level of expertise in movie making or 3D animation, etc.). At first, the experimenter provided a short demonstration of the tool and its different features. Then, participants were equipped with the head-mounted display and immersed in a training virtual

environment. During the training session, which lasted 15 minutes, participants were guided through a short tutorial and were allowed to ask questions to the experimenter. After the training, participants were immersed in a new virtual environment. The virtual environment used in the experiment is the 3D animated scene from a public dataset obtained from [14].

The main experimental task was to reproduce a given storyboard (see Figure 13), which was always accessible through a 3D panel. Participants were asked to produce a video sequence that best matches the storyboard composing of 12 distinct shots with no time limit. The task required participants to place a set of cameras and edit the entire sequence (defining the duration of the shots for each camera). Two shots also required pan camera motions. The entire editing process was done using our tool and the final result was a short film of 79 seconds. At the end of the experiment, they were asked to fill the NASA-TLX questionnaire in order to measure the effort required to use the tool and provide subjective comments regarding the ease of use and the usefulness of the system. In total, the entire experiment (including the training) lasted one hour.

We could not provide any comparison with other VR editing tools as at the moment, only our tool provides a complete film crafting and editing pipeline in VR.

### 5.1.3 Results and Discussion

All participants except one were able to finish the task in less than 30 minutes, one novice participant required one hour to finish the task. The accompanying supplementary material<sup>2</sup> provides more examples of the results obtained by participants.

All participants agreed that the learning curve of the tool was fast and would require just a few days to perfectly master the tool. This feeling is supported by the results of the NASA-TLX which showed a very low level of frustration (Frustration Level: *Median* = 2; *IQR* = 1) and a low level of required effort (Effort: *Median* = 3.5; *IQR* = 2). These results show that our tool could be integrated into a film making pipeline without a strong impact in terms of training, i.e. ensuring a good ease of use.

Another key aspect is whether the tool enables to efficiently perform the required task. The results from the NASA-TLX showed that participants achieved a strong level of achievement (Overall Performance: *Median* = 8; *IQR* = 1). This result can be interpreted as that users were pleased with the final obtained result. When analyzing the comments from expert users, there is a unanimous and strongly positive feeling for the tool, “It allows to create an animation sequence really fast!”, “It can be useful to pre-visualize complex scenes” or It is extremely useful in terms of camera placement and

to obtain a global view of the final editing.” The 3D animators and film school students said they would love to use the tool for their projects whether it is for storyboarding, for previsualization or for creating a first draft to be refined with a more advanced and dedicated tool (i.e. Blender, Maya, ...). Yet, the comments from the expert users shows that the obtained result might still not be enough for professional film-making editing (which was not the goal), but they could integrate the tool in project/animation films, which shows the maturity of the tool. For example, one participant would have liked to have additional edition tools “The editing features are less developed, it would be more useful to have a 3-point editing tool”, and another participant complained about the camera jitter “.. the camera could be more stable” due the high tracking quality and tracking rates of the HTC Vive device. Again, the obtained results could be exported to professional 3D modeling and film editing tools for a higher quality result.

Regarding how demanding it was using the system, NASA-TLX ratings shows that the mental demand was relatively low (*Median* = 3; *IQR* = 1.75), the physical demand was low (*Median* = 2; *IQR* = 1.75) and the temporal demand was relatively low (*Median* = 3.5; *IQR* = 2). These results provide a first assessment of the comfort while using the tool, which does not show any particular issue. Finally, regarding simulation sickness, although we did not directly measure it, participants did not show any particular symptom.

As a side note, one interesting observation is that novice and expert users followed a different workflow. Experts performed the tasks sequentially, first setting up all cameras and then performing the editing. In contrast, novice users constantly switched between editing and camera placement mode, i.e. placing a camera, then adding it to the edit and moving on to the next shot. This shows the potential need to provide different adaptations of the user interface to better match user workflows.

## 5.2 Experiment 2: Creative uses

Since our purpose is not to replace existing modelling and animation tools but to offer a complementary approach, it was not relevant to perform direct task-to-task comparisons with criteria such as time, precision or quality. To complete the first experiment, we rather proposed to rely on aspects such as acceptance through the TAM’s construct: Technology Acceptance Model, a direct, pragmatic instrument to measure a given technology’s degree of acceptance [24], that helped us to perceive the user’s satisfaction on an approach very novel for the practitioners.

We focused on the three TAM constructs:

- Perceived Usefulness (PU) – the extent to which the user expects the system to improve his job performance within an organizational setting;
- Perceived Ease of Use (PEU) – the degree to which the user believes the system’s use will be free of effort
- Attitude toward Using (AU) – the user’s desire to use or favorableness feelings towards using the system

### 5.2.1 Apparatus and Participants

We followed the same apparatus as the first experiment. Only the participants changed. In this experiment, we worked only with expert users, with at least five years of experience each in storyboarding, previsualization and 3D animation for the movie industry, or technical preparation of movie shooting. None of the experts in this session had participated in the previous experiment. Three experts were recruited. All of them are also teachers in film and animation schools.

<sup>2</sup><https://cinematography.inria.fr/vr-previsualisation/>

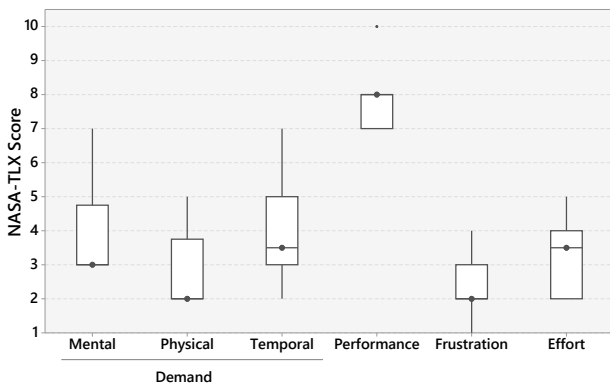


Figure 14: Boxplot of the results on the NASA-TLX questionnaire for the first experiment.



### 5.2.2 Experimental Protocol and Design

Upon their arrival, participants followed the same protocol (questionnaire, consent form, demographic information). Each expert was assigned to test one specific tool: the storyboard artist tested the storyboarding tool, the animator tested the previsualisation tool and the techvis artist worked with the technical tool. At first, the experimenter provided a short demonstration of the tool and its different features. Then, participants were equipped with the head-mounted display and immersed in a training virtual environment. During the training session, which lasted 15 minutes, participants were guided through a short tutorial and were allowed to ask questions to the experimenter. After the training, participants were immersed in a new virtual environment. The virtual environment used in the experiment is a basic reconstruction of a studio setup, with different types of camera rigs, light rigs and characters to animate.

During this stage of the experiment, there were no guidelines as to what to create; users had complete freedom to explore the features of the tools, yet they each had to design a simple sequence. At the end of the experiment, they were asked to fill a questionnaire with limited choices and open answers to evaluate the ease of use and usefulness of the tool. In total, the entire experiment (including the training) lasted one hour per users.

### 5.2.3 Results and Discussion

All users answered the same questionnaire. For the ease of comprehension, we will name them respectively the Storyboard expert, the Animation expert, and the Techvis expert. First, all three experts were satisfied by the results they designed – they evaluated their satisfaction at 8/10 for the storyboarder and 7/10 for the other two. Second, in evaluating the possible gain in terms of time, users were positive with feedback such as (Animation expert) “*the tool would enable the faster design of pitchvis previs and techviz sequences*” (pitchvis is the task of creating an animated sequence of the movie to convince producers to invest in the movie production).

Users were more dubious on the gain on the estimation quality of the generated contents, commenting that the tool would need an improved UI to better guide the users in their tasks. It was however highlighted that for the purpose of pitchvis and previs, such quality might be sufficient. On the Perceived Usefulness (PU) regarding the storyboarding, the Storyboard expert commented that while the tool would not provide him as much flexibility as a simple drawing, it could prove extremely useful for directors, producers and anyone involved in the creation process that are not used to sketch ideas on paper – a vast majority of them apparently.

The Perceived Ease of Use (PEU) was overall quite good (evaluated through questions such as “would it be easy to use the tool in current productions”), with the limit that more input/output geometry should be implemented to ensure a good degree of interoperability between 3D tools (exporting cameras only, exporting light sets, or Edit Decision Lists, etc). The experts estimated it would only take them between a few hours to a couple of days to master the tool.

The Storyboard expert commented that intentions and emotions were definitely more easily conveyed by drawings, but that this tool could help in preparing actions or dialogue sequences, where setups may be complex and suffer many iterations. One of the benefits listed by the Animation expert was that “the tool provided a better sense of space” than traditional 3D animation packages, hence giving more time to focus on the creative tasks than the manipulation of the camera perspectives.

On the Attitude towards Using (AU), the overall feedback of these experiments was positive, and all the experts agreed on the potential of such system for production purposes and were ready to use it in their production pipelines if technical aspects such as data interoperability were addressed, as well as more precise control of the animations. As teachers, they also appreciated its potential as an educational tool.

## 6 DISCUSSION AND CONCLUSION

With the growing maturity of VR technology, the presence of VR will grow in the professional market. This paper brings for the first time the use of VR for film storyboarding, previsualization and technical preparation by providing a tool that eases content creation for the different stages of film pre-production. The main contributions of the paper are the delimitation of the design space for VR storyboard tools, the iterative conception of the system and its evaluation. The interviews with experts have shown the requirement of letting creatives to keep control of the storyboarding process through simple tools. Thus, we have designed the system trading off expressiveness and control. For example, the editing step is simplified to enable fast prototyping at the expense of precise control. We would like to stress that the final result generated by the system could be further refined using other specific tools. Moreover, the tool could be used for content generation in videogames or VR applications. For example, it can be used to easily generate replays or “film” VR experiences.

Results demonstrate the benefits of the tool in rapidly prototyping filmic sequences for novice and professional users. Participants were able to achieve satisfying performances with just a few minutes of training and the results were overwhelming and extremely positive. All expert users were willing to use future versions of the systems and integrate the tool into their projects. In addition, the proposed technique demonstrates the ease with which tasks that are complex to perform with 3D modelling tools are made simpler by providing a seamless integration between different pre-production stages, and high-level interactions that take advantage of VR and 6DoF manipulation devices.

Finally, the user evaluation also helped to identify a number of limitations, such as the lack of stabilization for handheld camera motions and the lack of more complex film editing techniques. Future work will focus on the design and evaluation of more specific interaction techniques to provide higher expressiveness and control.

## REFERENCES

- [1] Five steps to avatar: reinventing moviemaking. [https://www.wired.com/2009/11/ff\\_avatar\\_5steps/](https://www.wired.com/2009/11/ff_avatar_5steps/).
- [2] FrameForge. <https://www.frameforge.com/>.
- [3] iClone. <https://www.reallusion.com/iclone/>.
- [4] MovieStorm. <http://www.moviestorm.co.uk/>.
- [5] Oculus. <https://www.oculus.com/story-studio/>.
- [6] ShotPro. <https://www.shotprofessional.com/>.
- [7] TiltBrush. <http://www.tiltbrush.com/>.
- [8] ToonBoom. <https://www.toonboom.com/>.
- [9] P. Baerlocher and R. Boulic. An inverse kinematics architecture enforcing an arbitrary number of strict priority levels. *The Visual Computer*, 20:402–417, 2004.
- [10] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 2004.
- [11] T. De Goussencourt, J. Dellac, and P. Bertolino. A game engine as a generic platform for real-time previs-on-set in cinema visual effects. In *International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS 2015)*. Catania, Italy, Oct. 2015. doi: 10.1007/978-3-319-25903-1\_76
- [12] D. Dwarakanath, C. Griwodz, and P. Halvorsen. Robustness of 3d point positions to camera baselines in markerless ar systems. In *Proceedings of the 7th International Conference on Multimedia Systems, MMSys '16*, pp. 21:1–21:12. ACM, New York, NY, USA, 2016. doi: 10.1145/2910017.2910611
- [13] Q. Galvane, I.-S. Lin, M. Christie, and T.-Y. Li. Immersive previs: Vr authoring for film previsualisation. In *ACM SIGGRAPH 2018 Studio, SIGGRAPH '18*, pp. 4:1–4:2. ACM, New York, NY, USA, 2018. doi: 10.1145/3214822.3214831
- [14] Q. Galvane, R. Ronfard, C. Lino, and M. Christie. Continuity Editing for 3D Animation. In *AAAI Conference on Artificial Intelligence*, pp. 753–761. AAAI Press, Austin, Texas, United States, Jan. 2015.

- [15] N. Hagbi, R. Grasset, O. Bergig, M. Billinghamurst, and J. El-Sana. In-place sketching for augmented reality games. *Computers in Entertainment*, 12(3):3:1–3:18, 2014. doi: 10.1145/2702109.2633419
- [16] J. Hart. *The Art of the Storyboard: A Filmmaker's Introduction*. Focal Press, 2007.
- [17] E. Heitz, J. Dupuy, S. Hill, and D. Neubelt. Real-time polygonal-light shading with linearly transformed cosines. *ACM Trans. Graph.*, 35(4):41:1–41:8, July 2016. doi: 10.1145/2897824.2925895
- [18] R. Henrikson, B. Araujo, F. Chevalier, K. Singh, and R. Balakrishnan. Multi-device storyboards for cinematic narratives in vr. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, pp. 787–796. ACM, New York, NY, USA, 2016. doi: 10.1145/2984511.2984539
- [19] R. Henrikson, B. De Araujo, F. Chevalier, K. Singh, and R. Balakrishnan. Storeboard: Sketching stereoscopic storyboards. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pp. 4587–4598. ACM, New York, NY, USA, 2016. doi: 10.1145/2858036.2858079
- [20] Y. Horiuchi, T. Inoue, and K.-i. Okada. Virtual stage linked with a physical miniature stage to support multiple users in planning theatrical productions. In *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces*, IUI '12, pp. 109–118. ACM, New York, NY, USA, 2012. doi: 10.1145/2166966.2166989
- [21] R. Ichikari, K. Kawano, A. Kimura, F. Shibata, and H. Tamura. Mixed reality pre-visualization and camera-work authoring in filmmaking. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '06, pp. 239–240. IEEE Computer Society, Washington, DC, USA, 2006. doi: 10.1109/ISMAR.2006.297823
- [22] L. Leite. Virtual marionette. In *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces*, IUI '12, pp. 363–366. ACM, New York, NY, USA, 2012. doi: 10.1145/2166966.2167049
- [23] C. Lino, M. Christie, R. Ranon, and W. Bares. The directors lens: an intelligent assistant for virtual cinematography. In *ACM Multimedia*. ACM Press, 2011.
- [24] M. G. Morris and A. Dillon. How user perceptions influence software use. *IEEE Softw.*, 14(4):58–65, July 1997. doi: 10.1109/52.595956
- [25] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd. Real time localization and 3d reconstruction. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1*, CVPR '06, pp. 363–370. IEEE Computer Society, Washington, DC, USA, 2006. doi: 10.1109/CVPR.2006.236
- [26] T. Muender, T. Fröhlich, and R. Malaka. Empowering creative people: Virtual reality for previsualization. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, pp. LBW630:1–LBW630:6. ACM, New York, NY, USA, 2018. doi: 10.1145/3170427.3188612
- [27] C. Nguyen, S. DiVerdi, A. Hertzmann, and F. Liu. Vremiere: In-headset virtual reality video editing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pp. 5428–5438. ACM, New York, NY, USA, 2017. doi: 10.1145/3025453.3025675
- [28] M. Shin, B.-s. Kim, and J. Park. Ar storyboard: an augmented reality based interactive storyboard authoring tool. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 198–199. IEEE Computer Society, 2005.
- [29] Q. Wu, P. Boulanger, M. Kazakevich, and R. Taylor. A real-time performance system for virtual theater. In *Proceedings of the 2010 ACM Workshop on Surreal Media and Virtual Cloning*, SMVC '10, pp. 3–8. ACM, New York, NY, USA, 2010. doi: 10.1145/1878083.1878087