EAI Endorsed Transactions on Creative Technologies

Research Article EAI.EU

Using Natural User Interfaces for Previsualization

Rainer Malaka, Tanja Döring*, Thomas Fröhlich*, Thomas Muender*, Georg Volkmar*, Dirk Wenig*, Nima Zargham*

Digital Media Lab, TZI, Bremen, Germany
malaka@uni-bremen.de

Abstract

INTRODUCTION: An important phase in the process of visual design for the narrative media is previsualization (previs). Professionals use complicated 3D software applications that are not especially designed for the purpose of previs which makes it difficult for the artists and non-technical users to create previs content.
OBJECTIVES: The aim is to empower artists to express and visualize their ideas and creative capabilities in an optimal way.
METHODS: We suggest using natural user interfaces (NUIs) and discuss suitable interactions for different previs tasks. We developed and evaluated a series of individual prototypes as well as a central overarching prototype following our NUI concepts.
RESULTS: The results show that our NUI-based interaction methods were perceived highly positive and experts found it valuable for their work.
CONCLUSION: With only a brief familiarization phase, NUIs can provide a convenient substitute to traditional design tools that require long training sessions.

Received on 04 March 2021; accepted on 15 March 2021; published on 16 March 2021

Keywords: Previsualization, Natural User Interface, Virtual Reality, Animation, Film, Theatre, Visual Effects

doi:10.4108/eai.16-3-2021.169030

1. Introduction

Previs is known as an indispensable and critical process that could increase the production value of a project [1, 2]. It has been adopted for decades in disciplines such as film, product design, and architecture and is growing in popularity because of technological improvements [3]. Most of today's previs task is handled digitally and despite all the potential advantages for production, it is time consuming and requires trained personnel. The tools used for the purpose of previs are frequently complex 3D tools such as Maya[1], Auto CAD[2], Blender[3] and recently game engines such as Unity[4] and Unreal[5] [4]. The use of such tools is time consuming and requires proficient technical experts excluding the non-technical, imaginative users namely artists, designers, and directors from the previs process. This requires various design iterations and constant communication between the technical experts and the creative staff [5]. This could be problematic due to the fact that most creatives are used to be in control of the design process by drawing, writing, or sketching.

Current 3D software programs used by experts for previs purposes provide an extensive set of features and functionalities, many of which are for highly specific use cases that are not really essential for previs. The collection of all these features within a 2D interface makes these tools complex and could frustrate users that intend to use them only for the purpose previs. Furthermore, several specific functions required for previs may not be supported by standard tools and a great deal of extra effort might be needed to find workarounds.

In order to design a previs software that particularly provides the necessary functions for the previs process,

*Authors are listed alphabetically
[1]https://www.autodesk.com/products/maya
[2]https://www.autodesk.com/products/autocad
[3]https://www.blender.org/
[4]https://unity3d.com/
[5]https://www.unrealengine.com

investigating the tasks that are specifically carried out in digital previs is crucial. In a survey study amongst practitioners [5], we identified and analyzed a persistent collection of previs tasks and the main functionalities that a reliable previs software should support. Previs applications user base also comprises of a broad range of users in different roles who have different workflows and stories: what is wanted for animation is not necessarily what is wanted for theatre [5].

We believe that users should have the possibility to choose their interaction method to use for different previs tasks and select the one best fitted for their creative needs. This way, the artist can express their creativity and be more focused on their previs task without having to face an inflexible interface that does not specifically consider the needs of an artist. To involve non-technical people more into the previs cycle and empower artists to use a digital previs software, we propose using natural user interfaces (NUIs) for previs tasks to empower users in how they can express themselves in an optimal way depending on the task they want to perform in the previs cycle. NUIs have the potential to minimize the effort needed to translate user's actions and provide increased direct natural expressiveness [6].

In the pages that follow, we examine the potential NUIs to support intuitive interfaces for the purpose of previs. Moreover, we provide a comprehensive analysis of digital previs in practice within the application domains of film, animation and theatre. We then propose recommendations to support such tasks with NUIs and create the natural interaction concept by mapping both users' needs and the previs task characteristics to meaningful and natural interaction techniques. In order to examine our interaction concept, we developed and evaluated a series of individual prototypes which we refer to as 'vertical prototypes', as well as a central overarching prototype which we call the 'horizontal prototype'. The prototypes were designed within the framework of the EU-project "first.stage". First.stage is a previsualization project that researches and designs natural user interfaces and tools for previs in order to make it accessible to non-technical users.

Using these prototypes, we aim to find how to integrate NUIs in the previs process of the creative personnel and how NUIs affect the production process in comparison to common 3D software.

The main contribution of this work is to provide a natural user interface concept that is built around a set of interaction methods which are chosen specifically to the nature of the diverse previs tasks. The concept is based on the requirements of the application areas and incorporates a detailed review of the literature.

2. Related Work

2.1. Natural User Interfaces

The term natural user interfaces refers to a broad collection of interactive technologies that offers rich ways for interacting with the digital world using existing human capabilities for communication and human capacity to manipulate the physical world [7]. Although it still lacks a common definition since the understanding of this term is subjective to the constant development of the technology [8]. However, in the literature, there are recurring aspects of NUIs that appear to establish a shared view on natural user interfaces. NUIs utilize the capabilities of human beings to express themselves with their body movement, voice and gestures, thus described as intuitive [9] [10]. The term NUI encapsulates any technology that allows users to interact in a more natural or humanistic way with computer systems [11].

Early research on NUIs roots back to the 1980s, where they introduced the "Put that There" system, an interface controlled by voice and gesture [12]. In the beginning, the term NUI was mainly used for (multi-)touch gestural interaction. Eventually, researchers and designers worked with different modalities such as speech, gesture3D, manipulation, and virtual and augmented reality (AR), expanding the terminology. Some popular consumer products that have been mentioned a lot with reference to natural user interfaces are the Wii[6] and the Kinect for Microsoft Xbox[7], both vision-based devices that indirectly or directly support full-body movements for system input [13]. The premise of NUIs is to let computers understand the innate human means of interaction and not induce humans to train to the language of computers [14]. They aim to provide a smooth user experience where the technology is almost invisible [15] and make users learn the interactions as quickly as possible. Liu describes the characteristics of NUIs as being user-centred, multi-channel, inexact, high bandwidth, voice-based, image based, and behaviour based [16].

A very similar concept to natural user interfaces is Reality-based Interaction (RBI) by Jacob et al. [17]. They proposed RBI as a unifying concept that ties together a large subset of the evolving human-computer interaction techniques. The RBI framework allows to discuss aspects of the multitude of current user interfaces "beyond the desktop" in a structured way by identifying and analyzing aspects of reality and computational power that are useful for interaction. The authors argue that the new interaction styles make use of users' pre-existing knowledge of the everyday,

[6] http://wii.com/
[7] https://www.xbox.com

non-digital world. Their assumption is that "basing interaction on pre-existing real world knowledge and skills may reduce the mental effort required to operate a system because users already possess the skills needed" [17].

Although you can see the use of humans' experiences from the physical world to a stronger degree in most recent user interfaces, computer tools and applications would not be as powerful, if systems were based purely on reality-based themes. The "computational powers" of computers such as Expressive Power, Efficiency, Versatility, Ergonomics, Accessibility and Practicality are undoubtedly advantageous [17]. That makes it challenging for the current user interface design to combine and find a balance between the computational power and reality. Jacob et al. propose that "the goal is to give up reality only explicitly and only in return for other desired qualities" [17].

2.2. NUIs for Performance Animation and Modelling

The common animation systems are too complex for the high demand in animated content today. To improve the accessibility and efficiency of such systems, researchers suggest an HCI perspective on computer animation [18]. An interaction perspective on computer animation can help to construct a design space of user interfaces for spatiotemporal media. Many researchers have investigated the use of NUIs for performance animation and modelling. Lee et al. presented the implementations of full-body input as a natural interface to character animation by extracting user silhouettes from camera images [19]. Chai & Hodgins [20] introduced an approach for performance animation based on optical marker tracking.

An affordable approach to performance animation is presented by Walther-Franks et al. [21] with the "Animation Loop Station", allowing users to create character animations layer by layer by capturing users' movements with Kinect sensors. In their proposed system, a speech interface is included so that users can fluently work on their animation without the need for a graphical interface. In another work by Walther-Franks et al. [22], authors introduce the Dragimation technique which allows users to control timing in performance-based animation on 2D touch interfaces where they can directly interact on the characters instead of a timeline. The authors found that Dragimation performs better with regard to learnability, ease of use, mental load, and overall preference compared to timeline scrubbing and a sketch-based approach. This system was inspired by the work of Moscovich et al. who introduced a rigid body deformation algorithm for multi-touch character animation [23].

An interesting approach to augment the own character animations with rich secondary animations is the combination of performance animation with physics simulation. However, the combination of both technologies is not trivial. An approach is presented in a natural user interface that combines motion capture using the Kinect sensor and physics simulation for character animation by Liu et al [24]. They introduce a framework that combines both technologies to prevent conflicting inputs from users' movements and physics engine. Their approach and framework has been extended by Shum and Ho (2012) who present a more flexible solution to the problem of combining physical and motion capture information [25].

Another area of complex 3D interaction is the field of digital modelling and sculpting. There is a strong need for tools that allow for natural expression in digital content creation, especially for previs, as many productions start with zero assets. This means, that assets, objects, and props have to be created from scratch most of the times. For instance, Herrlich et al. investigated interface metaphors for 3D modelling and virtual sculpting [26]. The first implementation of virtual sculpting was presented by Galyean and Hughes (1991) [27]. They used a custom force-feedback system that would translate the absolute positions of the input device into a 3D mesh. The approach was further extended by Chen and Sun (2002) and Galoppo et al.

by implementing virtual sculpting using a stylus device and a polygonal mesh instead of a voxel approach [28] [29]. Wesson and Wilkinson implement a more natural approach by using a Kinect sensor for deformation of a virtual mesh, while also integrating speech commands for a more fluent user experience [30]. Natural animation can also be approached by using Virtual Reality (VR) technology. For instance, Vogel et al. designed a VR system for animation where users work with a puppeteering metaphor for character animation [31]. They evaluated their tools with animation experts and found that it improves the speed of the workflow and fast idea implementation.

3. Previs Task Analysis

Typical previs tasks include scene layout, camera work, animation, and special effects [32]. To better develop a software that is specifically suitable for the purpose of previs and covering the fundamental functions needed for this process, we examined the tasks that are specifically operated in digital previs [5]. In order to do this, we first collected user requirements by interviewing experts from the domains of film (n=5), animation (n=3) and theater (n=8), using scenarios, personas, workflow generation, prioritization, and categorization, and extracted the everyday tasks and work-flows carried out in digital previs. After we gathered the requirements from the experts in the interviews, we arranged a workshop with the experts of the domain. We used the MoSCoW Analysis [33] to obtain the functional requirements which were not identified during the interview. We outlined a set consisting of the requirements which were collected within the interviews and the workshop. We produced index cards and employed color coding to mark the origin of every collected requirement for the reasons of clarity and organization.

Overall, we identified a set of 118 requirements. Based on the collected information, we identified 20 functionalities as core features that an efficient previs software should support. Although, one can not directly translate all these core features into previs tasks, since a few of these are general functions such as the ability to support multi-user and visual effects (VFX). Hence, we reduce the number of core functionalities to those that are essential for previs and can be applied to all application areas.

Result of this reduction presented the 9 core previs tasks: Sketching (Modelling), Project Structure, Shot Management, Import/Export, Camera Control, Assets and Layout, Animation, Lighting and Posing, and Visual Effects.

For further information regarding this procedure, please refer to Muender et. al [5].

4. Natural Interaction Concept

In this section, we present our Natural Interaction concept that builds the foundation for the implementation of our vertical and horizontal prototypes. We designed our core interaction techniques based on the core previs tasks that we identified, each one considering the natural aspects of the task and their context of use.

In Figure 1 we provide a graphical overview of our overall interaction approach based on the previs tasks mentioned earlier. We open up a 2D/3D space where we fit the task affordances to 2D and 3D interaction techniques and select the hardware correspondingly. Our ideas and concept development are driven by the variety of feedback on naturalness during the requirements elicitation that has been performed and our first-hand experiences with users that tested our prototypes. We further present additional interaction methods that complement our concept generation.

4.1. Interaction Techniques

It is crucial for a previs software to include the core functionalities of a previs task, "those functionalities of the product without which, the product is not useful for the users" [33]. Based on the requirement analysis, we suggest direct manipulation via touch on 2D interfaces, 3D direct manipulation, object-oriented user interfaces, spatially-aware displays, tangible interaction, AR, full body and embodied interaction, hand-based interaction, and speech as interaction techniques that would fit for a NUI based previs software. Here we discuss these techniques in more detail and show how they fit to the respective tasks.

Direct Manipulation via Touch. 2D interfaces that offer direct manipulation are intuitive and provide high accuracy. As with the high penetration of iOS and Android touch devices, these interfaces prove to be very intuitive for a large user base and naturally capture user intent on 2D screens via direct manipulation metaphors and 2D gestures. For tasks such as import/export, shot management, and project structure, a 2D touch interface is suitable and natural as it benefits from current mental model of users which they employ for 2D content manipulation.

3D Direct Manipulation. Direct manipulation is well suited for the interaction with 3D content as it is presented in a model-world interface which reflects to the real world. Having real-world metaphors for objects and actions can make it easier for a user to learn and use an interface. On top of that, rapid, incremental feedback allows a user to make fewer errors and complete tasks in less time, as they can see the results of an action before completing it. The user of a well-designed model-world interface can wilfully suspend belief that the objects depicted are artifacts of some program and can thereby directly engage with the world of the objects. This is the essence of the "first-person feeling" of direct engagement. This interaction technique is natural and well suited for all tasks that require object manipulation in 3D space using stereoscopic view, such as modelling, assets and layout, camera control and motion, animation and posing, and lighting. In practise, there can be situations when it is not possible for users to interact with certain objects directly. This is especially relevant for groups of objects which can be spread across the scene, objects that are far away from the user or are too small for direct interaction, and objects that are completely or partially occluded by other objects. In order to overcome such issues, we implement an addition to the direct manipulation concept: *surrogate objects*.

A surrogate object is a reification of one or more domain objects that the user intends to interact with [34]. Other than the object within the scene, a surrogate object can always be presented in the field of view of the user and within the interaction space, reachable by touch or VR controllers. The idea of surrogate objects shows that by extending the direct manipulation concept in such way, the limitations of direct manipulation can be overcome [34].

Object-oriented User Interfaces . State-of-the-art 3D tools often come with overloaded Windows, Icons, Menus and Pointer (WIMP) interfaces. Such interfaces are not suitable for typical users of previs software.

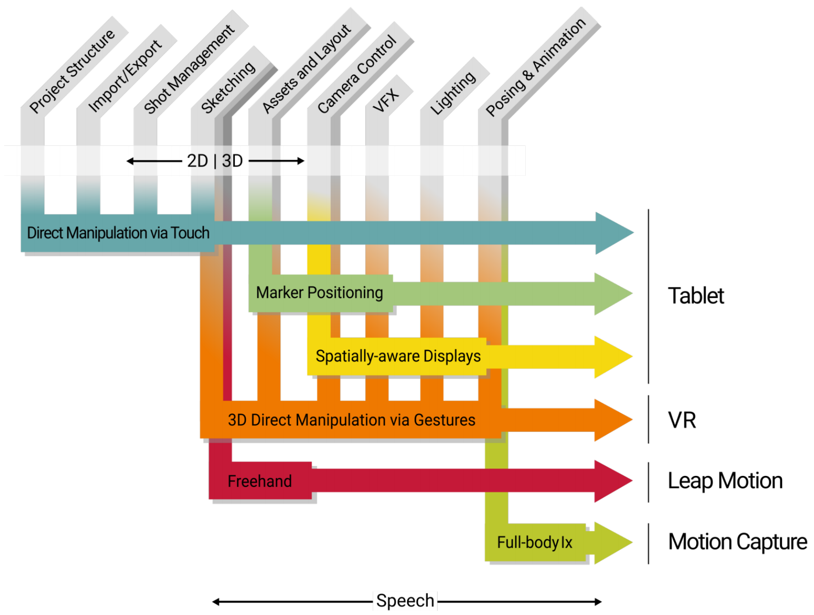

Figure 1. From previs tasks over interaction techniques to technologies

Previs tasks require specialized training and are often time consuming. As different input and output devices often require an entirely different interface, it may be required for the user to learn three different interface variants. These arguments contradict the requirement that the software is suitable for non-technical persons and can be learned quickly. A common practice in standard WIMP interfaces is to use buttons for all possible actions. If an object is selected, the actions that are not supported for this object are deactivated (in grey colour) but still presented to the user. This leads to cluttered interfaces, which can easily overwhelm users. In addition, buttons are commonly spread across the interface leading to a loss of association with the intended object that user wants to manipulate. Based on these observations, we implement object-oriented user interfaces (OOUI). These are types of user interfaces in which the user interacts explicitly with objects that represent entities in the domain of the application. They can be seen as the counter approach for function-oriented interfaces that are normally used in 3D applications. We propose a system in which objects have a dedicated interface only displaying the actions available for this specific object.

The interface is positioned relative to the object it is corresponding to, rather than fixed button positions in a static WIMP interface. Object oriented user interfaces are proven more user friendly compared to other interface paradigms and provide several advantages in terms of usability [35] . In addition, the relation to the object which will be affected by an action is better understandable. Users can also classify objects based on how they are presented and behave. In the context of what users are trying to do, all the user interface objects fit together into a coherent overall representation. OOUI can reduce the learning curve for new users as only relevant actions and options are displayed. It is possible to display an object related interface with a similar visual representation on different output devices such as tablet or VR. This can be rendered as part of the application and does not need a WIMP GUI. Furthermore, the interaction with such an interface can be designed to be similar with different input devices.

Spatially-aware Displays . These are displays that have information about their position and orientation in the room either by relative differences (gyroscope and compass data) or access to absolute position and rotation data using tracking devices. Using such displays, which are mainly tablet devices, one can create a direct access to a virtual scene by putting the control over the virtual camera directly in the hands of the users. They can then move this device across the room in order to change position and orientation of the virtual camera. This approach can be complemented with basic 2D gestures in order to extend the interaction space. Previs tasks that make the most out of this interaction technique are camera work and camera motion. Using spatially-aware displays, users can position cameras in a virtual scene and navigate in 3D while at the same time having the resulting shot visible through the display at all times for high efficiency. Thus, it is very natural to frame a shot in a virtual scene using real-world references by walking in a real room and orienting the device as needed.

Marker Positioning (Mobile AR) . With augmented reality, real images and projections of real scenes on screens such as mobile and tablet devices can be augmented with digital information. This bridges the real world with the virtual world, allowing for an immersive experience. Marker-based AR could be used for layout related tasks that require multiple users to work together in a virtual scene. The markers connect physical objects with digital media and enable interaction with the data via interaction with the objects. These interfaces are intuitive and easy to learn since users can directly observe the actions on the digital data which have direct mapping to their actions.

Full-body Interaction and Embodiment . Embodied interaction relies on the integration of interaction between humans and computers into the material and social environment. The recording of motion capture data allows for full body interaction and embodied performance animation through inertial sensors that can be worn in the form of a suit. This makes it easy and natural to record character animations since users can benefit from using their own body as input, directly transferring their motion data to virtual characters. In several previs cases and especially for character animation, this can make complex and expensive keyframe animation techniques superfluous. Such interaction techniques are natural as they do not rely on a translation between user intent and action. This approach has a low learning curve regarding the animation and almost no graphical interface is needed.

Hand-based Interaction . With hand-based interaction users can perform manipulations to digital objects using their own hands without using digital tools by tracking the hand and finger motion and applying the data for the manipulation of the virtual object. We employ hand-based interaction for rapid prototyping and modelling of 3D content and objects in the modelling and layout tasks. Here we focus on the workflow of modelling and layout where we support users in "getting their ideas quickly out of their head" by removing a graphical user interface (noUI), allowing them to concentrate on implementing their vision in either a 3D model or scene layout.

Speech . Speech commands are used as a supportive layer that can be integrated as a secondary interaction to complement a primary interaction. For instance, in layout, users can filter and select the asset library using speech commands. Generic commands include deletion, menu interaction, and basic manipulation tasks that are expressive by speech but hard to express through manual interaction, such as flipping an object, switching to a specific camera view, or switching between different views and zoom factors.

4.2. Interaction Principles

Having introduced our core interaction techniques, we now present interaction principles that enhance the interaction by implementing the notion of "how" the interaction in our tools should be. We organize these into general principles that should apply to all prototypes and specific principles that will only apply to certain, focused prototypes presented in this paper.

General Interaction Principles.

Task Context Depending on what users are doing, it is important to find natural interaction for each specific scenario. An interaction is natural depending on its context of use. Organizing a digital project such as files on Windows Explorer or Finder is more natural on a 2D interface as it stems from the digital domain of graphical user interfaces that most computer users are familiar with. As Jakob Nielson suggests, It may be difficult to control a 3D space with the interaction techniques that are currently in common use for 2D manipulation [8]. It would be very different when trying to organize, copying, duplicating or moving files in VR using other than 2D metaphors. An example for this is the "minority report vision" that has often been used to exemplify the use of gesture and 3D interfaces for desktop tasks. Studies have shown that these kinds of tasks are slower to perform in 3D and are cognitively more complex to achieve than with 2D GUIs [36]. We pick up on this notion and motivate a natural use depending on the context. For instance, arranging 3D objects in space as well as exploring spaces, getting sense of scale, and picking shots for cameras are most natural done in VR. On the other hand, project organization is best done on 2D interfaces. Transforming and working with digital content for organization is, as previously stated, most natural using GUIs. Another example is motion capture. Animating humanoid characters can be done in different ways. It can be done by key-framing a 3D character or by drawing frame by frame. Both has advantages and disadvantages, but looking for a natural way of interaction, using the own body is the only way of having a direct 1-to-1 relationship between user intend and desired outcome. Using the motion suit, users can work with their own body without having to understand complex key-framing or drawing techniques, making it more natural to create animation content and providing a low learning curve.

Easy to Use In order to make a previs software accessible to users with little technical knowledge,

[8]https://www.nngroup.com/articles/2d-is-better-than-3d/

special aspects should be considered to make the software easy to use. Interactions should not be designed to be as fast as possible but plausible and intuitive to the user. The system should provide feedback for the actions and help users to follow the intent. This is especially important when an action involves multiple sequential interactions. In order to not overwhelm the user with options, only a minimal amount of possible options should be shown which can be achieved by using interaction scaffolding and nested interactions.

Consistency The interactions should be consistent within the application. Ideally, users only have to learn a minimal set of interactions. It should be avoided to use different interaction schemes for the same tasks in a different context. For example, positioning an object by “grab and place” should work the same way in layout mode and in animation mode. The interaction should be as consistent as possible across different input devices. This will help the user to seamlessly switch between devices and not have to learn or remember special interactions for this input method. In some cases, this will not be possible as the different input devices provide different input modalities and degrees of freedom. But a primary interaction (left click, tap or trigger button) should perform the same action on all devices. Furthermore, the interaction should be consistent with other applications from the field, so that the user who has learned to interact with another tool does not get confused by a completely different interaction scheme. This means the application should not break with interaction standards from the field, e.g. support drag and drop.

Feedback The system should provide feedback for every action performed by the user. Visual feedback should be provided on all hardware platforms. If the object the user is interacting with is currently not in the field of view, and visual feedback is not visible, assisting indicators at the edges should indicate where the interaction is happening. Haptic feedback should be utilized when interacting in virtual reality using the vibration functions of the controllers. Audio feedback is also helpful in certain situations especially when something is happening outside the field of view.

Nested Interaction Rather than putting all actions onto different buttons, the system should utilize nested interactions which chains the selection of an action, parameter finding and performing the action into a series of small lightweight interactions. Sequenced actions should be designed so that experienced users can perform them very quickly and do not perceive them as impairing. The buttons in a sequenced action should be positioned in close distance, following the direction of motion.

Specific Interaction Principles.

Reality-based The interaction in 3D should orient towards the interaction with objects in the real world. As the interaction in 3D space is still novel and is not a known for many users, it should utilize their everyday knowledge [17]. This applies to positioning and rotating objects: Small objects like a bottle can be grabbed, rotated and placed with one hand. Bigger objects like a table or houses on the other hand have to be pushed and rotated using two hands. Falling objects stop when in contact with the ground or other objects. This behaviour should be supported as it is perceived as natural to the user. Another aspect that should be reality-based is the interaction with buttons, knobs and slider elements from the object-oriented interface. A clear 3D representation of the intractable object should be provided comparable to real life light switches, volume knobs and radio controls. The visual representation should present feedback of the current state of the control.

Playful and Fun Playing is intrinsically motivated and autotelic[37]. When users are provided with a user interface that supports playful expression, they can explore, be creative, try different solutions, and find joy and amusement even in productive contexts. This increases long-term motivation and lowers the frustration barrier [38] which is achieved by more pleasurable interaction rather than optimizing for speed and being goal oriented. Playful interaction invites users to discover features rather than frustrating them with an overloaded interface. Therefore, it is suitable to support novice users and users with little technical knowledge.

Creativity Support In order to support the creative process of the user creating a virtual set, animation or camera shot, the interaction with the software can be designed accordingly. The system should invite users to interact with objects in a natural manner rather than telling them what to do. This can be achieved using affordances [39]. Designing subtle affordances invites the user to discover through exploration. Presenting users with different viewpoints on a scene or with the sequence of their actions can support them in their iterative and evaluative process.

Rapidness, Accuracy and Precision The interaction used should be rapid, avoiding time-based interaction and large motions. Time based interaction disrupts the workflow, especially for experienced users. For instance, “look and hold” or “point and hold” interactions could easily frustrate users. In contrast to the general schema of rapid interaction, it should be possible to improve precision with additional effort. For instance, scaling up in order to work on detailed structures.

Multi-User Handing over a project to others is complicated and requires a lot of management.

However, working together in the same physical context where a common understanding is shared in the physical space through observation and communication of intentions and tasks is natural. This natural interaction could be achieved by providing input to the system that is not bound to one device. For instance, in a multi-user scenario, different devices can be used in a shared context in order to achieve a common goal.

5. Concept Implementation and Evaluation

In this section, we present research prototypes developed and evaluated during the project. These prototypes were built to evaluate individual interaction concepts and principals.

5.1. Virtual Reality for Previsualization

In order to support the non-technical professionals to work with 3D content and allow for more expressive interaction, we suggest using virtual reality for previs. In contrast to traditional 2D interfaces, VR offers immersion, illusion of *embodiment*, and illusion of physical interaction for a more natural interaction (*reality-based*). We developed two prototypes in VR and conducted a user study with non-technical domain experts to investigate how they experience VR use for the purpose of previs.

Our first prototype was a tool that focused on camera work for taking shots. The second one was meant for exploring stage designs and experimenting with stage equipment in a theater. These two scenarios were chosen as they interest professionals from all relevant domains such as film, animation, and theater by covering important aspects of previs such as exploring a virtual scene, layout, and shot finding.

Our results showed that the participants were predominantly positive towards VR for previs and rated it as a useful technology. All users could perform typical 3D previs tasks even after a very short learning time. Using the prototypes also convinced participants of the practical use of VR for previs and removed certain doubts [3].

Figure 2. Camera prototype with yellow virtual camera (right). View finder is attached to left controller (left).

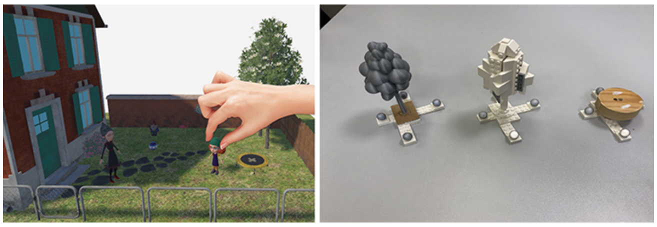

Figure 3. Left: The tangible scene design concept. Right: The 3D-printed tangibles with the highest fidelity

5.2. Tangible Scene Design

Using physical objects and models to visualize scenes is a common practice among professionals. Technologies such as VR have the ability to provide immersive experiences with an accurate dept and scale perception. Nonetheless, such technologies usually lack the tangibility aspect. In this work, we combined the VR technology with tangibles to utilize the immersive experience with the intuitive controls of miniature models, following the concepts of *3D direct manipulation* and *full-body interaction and embodiment*. We aimed at providing an interface where the non-technical users can intuitively create previs content (*creativity support*) using natural forms of interactions to create and delve into the 3D scenes. The prototype was used to evaluate how the tangibles with distinct haptic fidelity can effect the performance, immersion, and intuitive interaction for creating a 3D scene in VR. In a user study ($N = 24$), as well as an interview with eight experts in previs, tangibles with distinct production processes were compared (uniform objects without resemblance to digital object, 3D-printed and Lego-build tangibles similar to the digital object).

Our results showed that Lego was preferred as it provided *fast* assembly and adequate fidelity. No significant differences were observed in terms of haptics, grasping precision and the perceived performance. Professionals pointed out that digitally planning and adopting scenes is major benefit that could save a lot of time (*rapidness*) [40].

5.3. Collaborative Scene Design across Virtual Reality and Tablet Devices

Through the use of head-mounted displays (HMD's) users can perceive 3D content with believable depth sensation, being able to walk around and look around virtual objects. Further, room-scale tracking allows for natural movement (in a restricted area) and object manipulation in a natural and direct way.

VR controllers are used as extensions of the own body, employing easy and intuitive ways interaction by grabbing, holding, rotating and placing objects in VR. However, there are aspects in VR that limits the usefulness of the approach. Specifically in collaborative design, where several users work together in achieving a common design task, VR may restrict the collaboration. For instance, in VR users are often isolated and are not at the same location with other users. To share a common working space and work together, they are commonly connected with other users in VR. This interaction is usually done remotely rather than a shared office space as most places do not host more than one or two VR sets. This may have effects on the collaboration and work distribution between users.

In order to address these concerns, we aimed at combining the advantages of VR and tablet interaction to build a system where users can collaborate using both devices in a shared environment utilizing *direct manipulation via touch* and *3D direct manipulation*. We built a prototype where two users can work together simultaneously (*multi-user*) using VR and a tablet device in an attempt to improve work efficiency and *rapidness* (see fig 4). In a lab-study with 18 participants (9 teams) we investigated the impact of the device-dependent interactions on user behavior. Results showed that there are device-dependent differences in the interaction style that also influence user behavior. For instance, tablets were primarily used for overview and rough positioning, while VR was mainly used for smaller object manipulation. However, we did not observe a device-dependent "lead" role as the tasks were mainly distributed and worked on in parallel [41].

Figure 4. Witch house scene from the tale Hansel and Gretel created during the study

5.4. Tablet Camera Prototype

Following the concept of *spatially aware displays*, the tablet camera prototype provides the view through a virtual camera into a scene. The camera can be controlled by physically moving and turning the tablet. In addition, the camera can be moved forward and backward in view-direction by a slider. The camera motion can be recorded and replayed. This prototype is intended to give a more natural feeling (*reality-based*) while controlling virtual cameras by giving the user the tablet a physical device as a viewfinder for the camera. The prototype was evaluated by three experts in the context of film productions. The participants got an introduction to the tool and afterwards got a task to record a defined camera motion. In the end they were free to continue using the tool on their own. Semi-structured interviews were conducted with the participants to find out about their experience using the tool.

The results showed that the tool was perceived highly positive by the experts and they found it intuitive as they are used to having something physical in their hands as a viewfinder. All participants were able to record the defined camera motion in under five minutes.

Figure 5. The tablet camera prototype

5.5. Embodied Interaction for Animation Capture

Another important part of previs is character animation. Only by providing believable and interesting animations, previs becomes a convincing tool for production planning. However, the creation of character animation is a task for expert animators using complex tools that requires years of experience. Motion capturing could be a possible solution for such problems. However, there are still several issues with motion capture. One of the biggest is that individuals who work alone have problems with animating a scene completely without context. When wanting to create a scene with multiple people interacting with each other, different animations from each character has to be recorded with the other movements in mind so that the complete scene makes sense after editing. In order to overcome this issue, we developed a VR user interface for capturing embodied animations. *Embodied interactions* resemble bodily experiences that every human is familiar with (*easy to use*, *reality-beased*). In contrast to traditional interfaces for motion capture, this system enables users to record animations from the perspective of their own body, to slip in any other body (human or not) and perform animations from their perspective (see fig 6).

We performed a preliminary user study with 16 participants. Most users quickly became familiar with the VR user interface and interaction styles. We observed that even older participants could sufficiently work in VR after having a short familiarization phase. Nonetheless, we observed the need for another interface when it comes to integrating the tools in a larger working context.

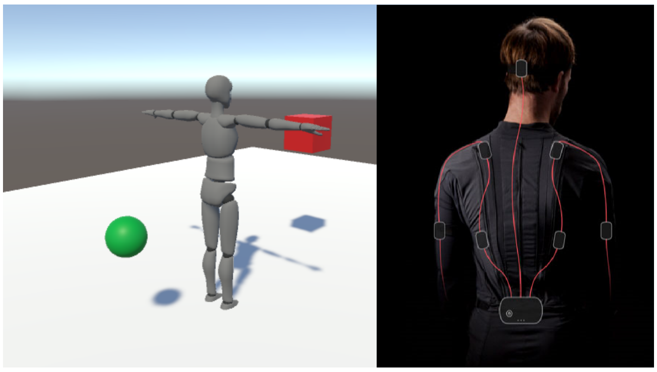

Figure 6. Users had to perform two basic and two advanced tasks using a motion capture suit. Basic tasks included interaction with the two objects, a green sphere and a red box.

5.6. Anticipation Cues for Motion Capture

To further develop the *embodied interaction* for animation capture, we integrated anticipation cues in the form of target indicators, progress bars, and audio feedback to help the user to record animations in sync with previously recorded animations or other animations in the scene. To design a prototype integrated with the anticipation cues, we interviewed three experts from the animation domain. These cues offer multi-sensory indicators for where and when to start the next animation. In a first step, the anticipation cues can be set by the user at the desired time and location. When going into the record mode, the cues are displayed in the field of view of the user, at the target location and are audible as sounds with increasing frequency as well as through the vibration of the controller.

The prototype was evaluated with 20 students to identify differences between multi-sensory anticipation cues and only visual cues. The results showed a clear preference for multi-sensory anticipation cues. Participants stated that they would not have been able to record the more complex animations without the cues.

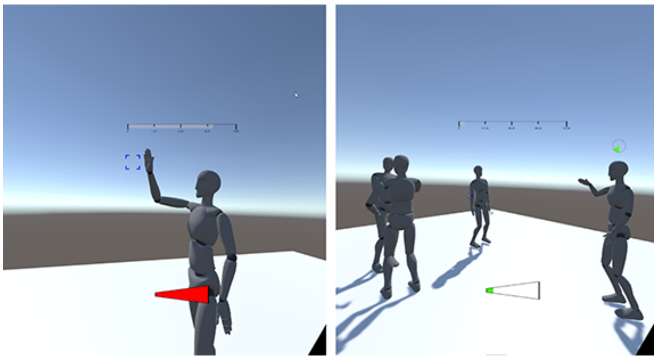

Figure 7. Anticipation cues: When recording an animation, the cues are displayed in the field of view of the user at the target location.

5.7. Playful Shooter for Fast and Easy Scene Design

Generating self-made content, e.g. for video games, allows players to creatively extend their gaming experience with additional maps or mods. Besides the player enjoyment of the additional content, production studios can benefit from user generated content since long-term motivation and playability is supported, at the same time production costs for studios are minimal. For this, games that support user content are often accompanied with tools for map or mod making like Valve's Hammer Editor [9]. However, these world or map builders can be cumbersome to use because of complex user interfaces with a high learning curve and limited UX.

In order to overcome this, we introduce a world builder game where the process of creating a 3D scene is play itself. Users of are able to navigate directly in the 3D scene, placing objects in a *playful* manner by shooting them in the scene for placement and manipulation. We implemented four weapons: A gun for physics enabled placement, a laser gun for more precise positioning and manipulation, a sniper rifle for far distance interaction, and a hand grenade for spawning multiple objects at the same time using an explosion, all allowing for playful interaction with the 3D content. We assumed that users would find the design of 3D scenes *easier* and more *fun*, feel more *creative*, have less difficulty positioning objects in 3D, and have a lower learning curve compared to standard editors.

We conducted an experiment with 17 participants where we compared our game to a simplified version of

[9]https://developer.valvesoftware.com/wiki/Valve_Hammer_Editor

the Unity editor, representing usability and workflow of standard 3D tools like Hammer. We were interested in how users perform both in a free building task and in a replication task where they had to rebuild a 3D scene from a printed template. Results show that users would like to use the game more frequently, found it less complex, would not need technical support and appreciate the learnability and ease of use. Additionally, we found that hedonistic and, interestingly, also the pragmatic quality of the game is rated higher than for the editor. This came as a surprise since the editor is highly task oriented and pragmatic in its design, functionality, and user experience [42].

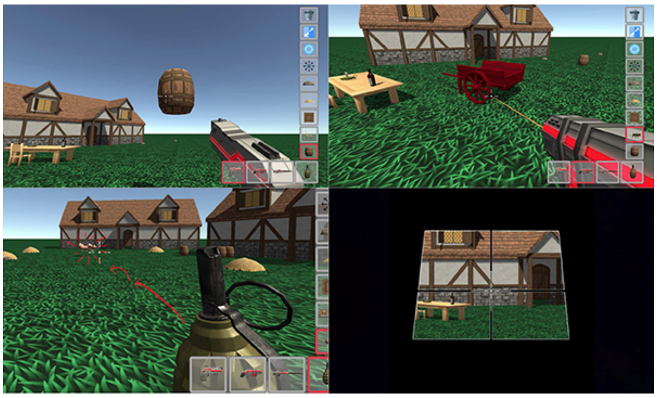

Figure 8. Overview of the implemented guns (clockwise): pistol, laser, sniper, grenade.

5.8. Rapid No-UI Modeling in VR

Creating 3D content, especially the aspect of modeling in 3D can be a challenging task for novice users. Available applications such as Blender or Maya are complex tools that have a high learning curve for beginners. As much as these applications are ideal for their purpose, that is the creation of complex models including shaders, lighting, animation, etc., beginners can feel overwhelmed by the user interface and the manifold options.

To address these problems, we looked into how to make the creation of 3D models and prototypes more natural and approachable for novice users in order to create a pleasurable experience that motivates diving deeper into the world of 3D modeling. Rather than creating detailed and complex models, the focus was on creating 3D models that represent a rough version of models that users have in mind through offering *rapid prototyping* with natural ways of interaction. Therefore we explored the use of *hand-based interaction* in combination with drawing gestures. Different mid-air gestures can be used to create primitive objects, e.g. cubes, spheres, pyramids. These objects can be grabbed through *direct manipulation* and assembled to more complex structures.

The prototype was evaluated by eight participants through structured interviews. The results show that the prototype was successfully used to create a variety of scenes in short time. However, we could also show that gestures that are too similar quite negatively influenced the mental workload and the users experience in general due to higher rates of miss-detection. However, this was mainly due to inaccurate and unstable tracking and the qualitative feedback was promising.

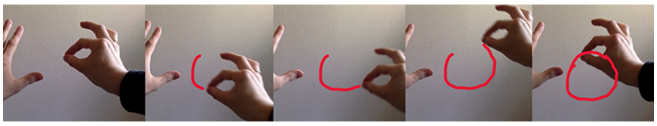

Figure 9. Demonstration of drawing a sphere gesture in mid-air using the starting gesture

5.9. Playful VR Sandbox for Creativity and Exploration

Sand playing is a nostalgic activity that reminds one about the childhood days where it opened possibilities for observations, creativity and cooperation. Sand play provides different developmental benefits such as proprioceptive sensing, body awareness, space awareness, and social skills. Such pleasant activity is not only limited to children, but many adults also enjoy sand playing and rediscovering the sensations of this activity. Interactive and augmented sandboxes take this experience to a new level by adding another interactive element to it, increasing the interaction and application possibilities. Simultaneously, people can play with the sand and experience projected interactive visual feedback onto it. We developed VRBox which is an interactive sandbox for *playful* and immersive terraforming. Our prototype combines the approach of augmented sandboxes with modern Virtual Reality technology and mid-air gestures. By exploiting *3D direct manipulation* and *embodiement*, this system extends the current approaches by adding a new interaction and visualization layer to interact with the sand and virtual objects. Furthermore, to provide better *creativity support*, users have the possibility to switch between two perspectives, a tabletop mode for the creation phase, and a first person mode by teleporting directly into their own creation which allows for a more immerse experience of the landscape in full-size. To evaluate the VRBox, a qualitative user study was conducted with nine experts from the domains of education, computer graphics, and game design. Our focus was on the user experience, as well as the technical aspects and the possible use cases. Our results indicated highly positive attitudes towards the VRBox which highlights the immersive and creative experience the system offers [43].

Figure 10. The landscape can be shaped by interacting with real sand (Right). VRBox from first person perspective (Left).

5.10. Grasping Objects in VR

With this prototype, we propose emulating human grip abilities for virtual reality to enable an interaction with virtual objects (*3D Direct Manipulation*) that corresponds better to object manipulation in *reality*. We devised different interaction designs to allow the user to dynamically set the firmness of the grip and thus be able to hold an object firmly and loosely using conventional controllers.

A qualitative pilot study and a quantitative main study have been conducted to evaluate the various interaction modes and to compare it to the current status quo of invariable grip. Pivotal design properties were identified and evaluated in a qualitative pilot study. Many test persons appreciated the suggested interaction's similarity to real object handling. The study participants especially used and valued the variability of their grip in vertical tasks in which the angle between object and hand typically needs to be altered. Two revised interaction designs with variable grip were compared to the status quo of invariable grip in a quantitative study. The users performed placing actions with all interaction modes.

Results showed that the variable grip has the potential to *enhance usability*, improve *realism*, reduce frustration, and better approximate real life behavior depending on tasks, goals, and user preference. The research conducted within the scope of this work contributed valuable insights regarding the manipulation of objects in a VR environment. To be precise, we gained an understanding of how to enrich objects with realistic rotational features in order to enhance overall usability and decrease mental effort as much as possible [44].

5.11. Drawing for Asset Selection

This prototype examines the feasibility of novel 3D object retrieval user interfaces in the context of VR

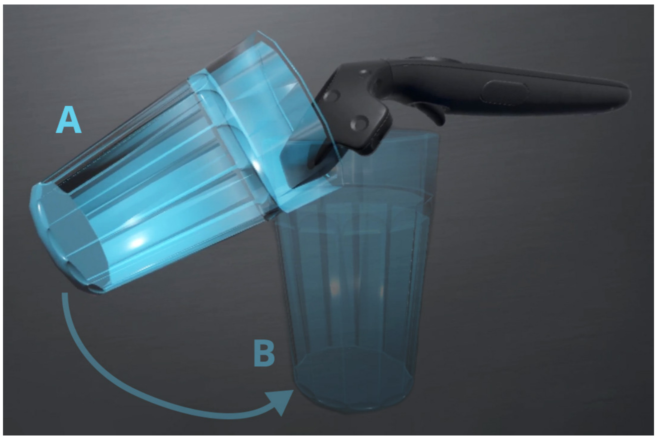

Figure 11. Illustration of grip adjustment. The glass is held with firm grip (a). Then, the user loosens the grip by releasing the trigger half-way: The glass swings downwards according to gravity (b). The position of the controller does not change.

and previs. Three prototypes were developed: sketch interaction, virtual keyboard interaction and *speech* interaction. All prototypes shared a common back-end in which assets were stored and associated with tags in a one-to-many configuration. Every retrieval type maps user input to asset tags.

In a within-subjects design user study (n=15), we examined the usability of each prototype. The results showed that applications of multimodal queries in an asset retrieval environment are promising. Overall, exploring the design space in a more natural direction proved more successful with the Voice interface. Though not in all cases statistically significant, its usability metrics fell within closer reach of the Keyboard interface. Moreover, the Keyboard interface exhibited significantly less task load. A rudimentary object placement system was devised in order to enable testing with complete tasks found in previs workflows.

5.12. Natural Language for Scene Interaction

This prototype investigates the possibility to arrange objects in a virtual scene using *speech* in an attempt to provide *fast* and *easy to use* interaction. It focuses on understanding relative spatial relations and ambiguous spatial descriptions, e.g., "next to". The prototype was developed using the speech API Wit.ai [10] to process the audio and to identify spatial keywords. A model of people common understanding of ambiguous spatial descriptions was developed for to translate the words into executable actions for the system. This model was developed based on a user study where participants had to place object according to ambiguous spatial

[10]https://wit.ai/

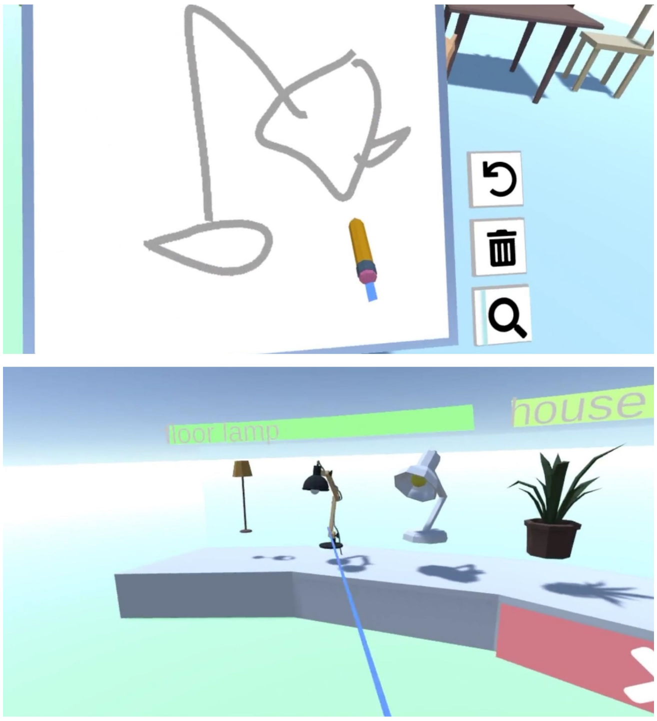

Figure 12. Top: The user draws a sketch of the lamp. Bottom: Based on the sketch the user can choose between different 3D objects.

descriptions. The data of this study was used to mathematically define most probable interpretations of the descriptions. The resulting model is used to translate spatial keywords from user input into positional data which can be used to position the objects.

The developed prototype was evaluated in a final user study with ten participants. The results showed that the task-load of the system is relatively low and in the same range as other input modalities such as mouse or touch. The participants could successfully position the objects of a scene according to the tasks they were given. Nevertheless, applying this approach to general scenes is challenging as all objects have to be annotated in order to define their up, front, etc.

5.13. Cognitively Enabled Scene Design

This prototype investigates a way to combine the power of modern Artificial Intelligence (AI) with the demands of professional designers to fine-tune their content. It incorporates a novel concept called Cognitively Enabled User Interfaces (CUI). These interfaces provide an innovative and *easy to use* interaction to *support creatives* by taking the user's cognitive world into account. To give an example, instead of adjusting the appearance of a 3D scene by tinkering with various menus and options, a scene designer might simply want to make the environment look more "exciting". A CUI would be capable of understanding this intention and have an underlying AI that knows what parameters need to be changed in order to get the desired outcome. For this purpose, we imitated an AI with the option to adjust a scene between the states of "cozy" and "scary" and integrated this approach into a simple slider interface. As a first evaluation of this approach, we implemented a prototype and conducted a comparative user study (n=31). We found that CUIs can offer a significantly higher usability and better user experience than traditional interfaces from the domain of virtual content creation.

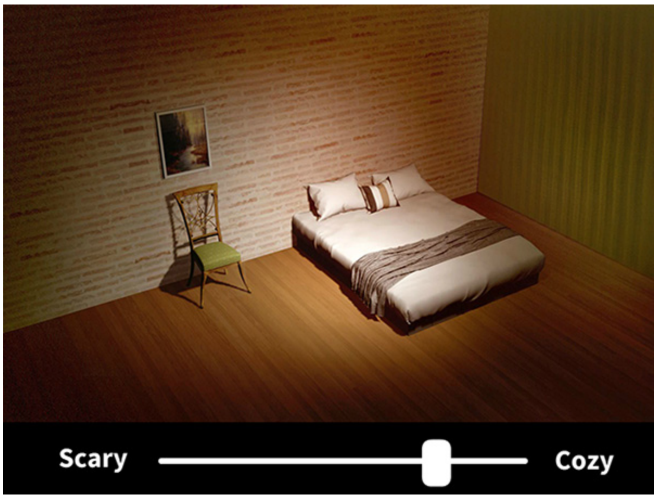

Figure 13. Example of a CUI capable of switching the scene between the states "cozy" and "scary"

5.14. ScenARy: Augmented Reality Scene Design

This *mobile AR* prototype is an application which can be used to layout scene by augmenting a table or room with the virtual objects of the scene. Virtual objects can be chosen from an asset library or created by the user drawing on the screen. In this *marker-based* prototype, regular playing cards are recognised by the application and virtual objects can be assigned to the cards. This enables the user to arrange the virtual objects by interacting with the tangible cards instead of the screen that displays the scene. The prototype was evaluated with professionals from the theater domain. After an introduction to the application, each participant had to perform a defined task using the app.

The professionals rated the prototype with a high usability score. The interview results demonstrate that this prototype can be a valuable tool for the initial planning phase of stage productions. The participants mentioned that they liked being able to physically layout the scene using the playing cards and did not have to work with complicated on screen controls.

The selection of objects to assign to the cards was perceived as textiteasy. All participants mentioned that is valuable that the tool encourages collaboration and *multiple persons* can use it together as the layout phase is a collaborative process. The professionals also mentioned that they miss some features such as lighting in order to use it for their professional workflow.

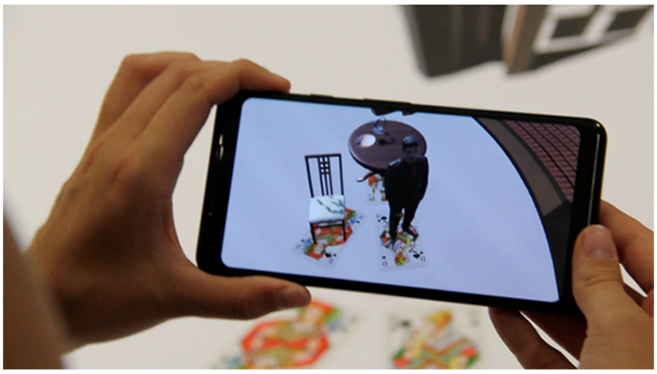

Figure 14. The AR scene design prototype running on a smartphone. The cards presents assets.

6. Main Prototype

Based on different interaction metaphors which are derived from our NUI Concept, in order to best support the various tasks and diverse users of a previs software, We developed a fully integrated software prototype which provides functionality for all relevant previs tasks defined earlier, e.g. import, sketching/modelling, set layout, lighting, animation, camerawork, motion capture and visual effects. We integrated the individual research prototypes or parts of them into the main prototype based on positive evaluation, technical feasibility and positive review by the application partners, in order to create a direct impact of the research results.

This prototype is a VR based application that provides a common space for performing previs tasks: the user can move freely within the virtual space and *directly manipulate objects in 3D* using the VR hand controllers. It provides the basis for a natural user interface where assets can be grabbed and moved directly, characters can be posed by grabbing and moving parts of their body, paths can be manipulated by grabbing nodes, and special tools can be wielded to perform tasks such as sketching and painting. Additional controls are provided by a tablet-like interface on the back of the user's wrist. *Touch gestures* and *speech* commands are also used to augment the interface. To support realistic work sessions involving many tasks as part of a larger workflow, the basic platform also provides an online asset repository and a shared repository for saving and reloading stage design and scene scripts along with undo and redo and *multi-user online collaboration* between team members. Scenes are organized into projects and each project has its own team. Theatre, animation, film, and visual effects each have different requirements from their respective user stories. This prototype provides a common core of functionality across all disciplines as well as specifically targeted functionalities for individual domains.

Figure 15. Possible previs outcome: 3D representation of a theater stage with different lights, props and virtual actors.

6.1. Scene Design

A fundamental part of previs is set design and asset layout. One can start with an empty stage and start building on it or simply import a model of an environment. The core functions for dressing the set and laying out assets are searching the repository for assets and placing them within the scene, precisely translate, rotate or scale the asset, selecting multiple objects to be moved or deleted as a group and attaching objects to others as parent/child such as cups on a table or lights on a rig. For the selection of the assets, on top of the primary mode of selecting an asset using the VR controllers, a *speech* interface is developed to retrieve the spoken objects. The object placement is carried out using *direct manipulation*.

In addition to regular static assets like a table, special assets that provide some functionality are contained in the repository and can be placed like any other object. These special objects are lights, visual effects, characters and cameras.

For **lighting**, several different styles of light are included such as a light with visible beams. Once placed in the set, users can adjust the color, brightness, range, and opening angle of the lighting.

Visual Effects provide dynamic objects which can represent fire, water, smoke and more. They are based on particle systems which can be activated and looped by the user.

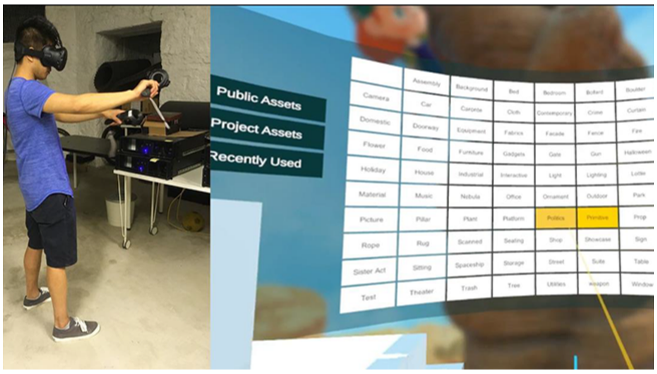

Figure 16. To place assets in the scene, the artist can pick an item from the asset library.

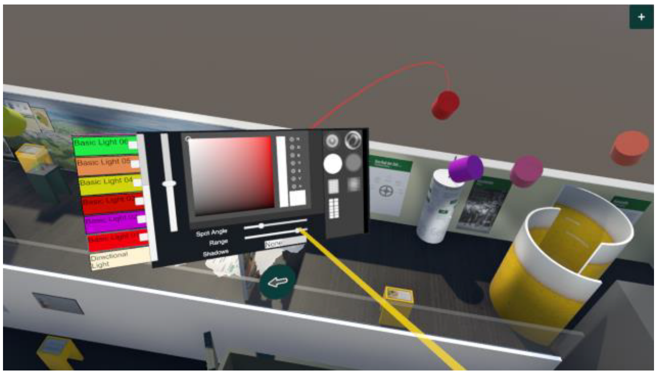

Figure 17. Light settings can be adjusted in terms of color, intensity, spot angle and other options.

Characters are most typically humanoid puppets but also include vehicles. These assets can be animated in detail and are the foreground of previs and story-telling.

In addition to the assets from the repository, users can create their own objects through **Digital Crafting**. Using the a Sketch Tool, one can create a shape by drawing an outline of it and modify the shape by grabbing and moving faces, edges or vertices of the shape. Once a shape has been created, it can be painted with an airbrush tool.

6.2. Animation

A major part of the previs process is animating characters and other objects according to a storyboard. The main prototype provides different animation possibilities for characters and objects to enable this. All animations are recorded on a timeline which holds every activity that is happening in the scene.

Character Animation. One of the most common needs in animation is to have a walking animation. In this prototype, one can simply sweep out the desired path for the character to follow from one mark to another. This path can then be modified by grabbing and moving it. Humanoid characters will have automatic behaviors such as walk, run, jump, fall, get up etc. Vehicles will also follow paths automatically applying steering, acceleration and braking as required.

Figure 18. Humanoid puppets which fall into the 'Characters' category.

Humanoid characters can also be animated in more detail through **posing**. Each character has handles attached to their body parts, e.g., shoulders, elbows and hands, allowing the puppet to be posed through *direct manipulation* (see figure 19). Poses set keyframes on the timeline and produce a path in space that allows the pose to adjust over time.

One of the easiest and most intuitive form of animating a character is through **Motion Capture**. The prototype supports the use of the Rokoko Smartsuit [11], allowing the user to directly take over the control of a character. Through the *embodiment* of the character the user can perform an animation in a natural way and is not restricted by complex controls(see figure 20).

Object Animation. Objects of a scene that do not fall in the category of characters can be animated through **rigid animation**. This provides a simple and intuitive way to animate by acting out the motion. A user can grab and move an object while recording the performed motion. This form of animation also refers to the concept of *embodiment* and *reality-based interaction* as it follows the idea of how kids play out scenes with toys. Once an animation path has been created, the path can be further manipulated by grabbing it at points and moving it as desired.

[11] https://www.rokoko.com/

Figure 19. Puppets can be posed by adjusting the attached handles to their body parts.

Figure 20. A performer directly animating the character using the Rokoko Smartsuit.

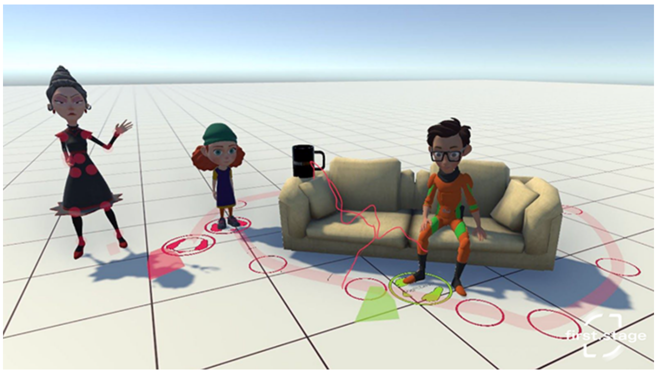

Figure 21. Rigid animation of a mug which can be manipulated by grabbing it at points and moving it as desired.

Objects in the scene can also be animated through **physics interaction**. This can be used as a quick and easy way to add natural secondary motions to assets. For instance, a chair might be knocked over as a character walks past it, or a character might kick a ball that then bounces off. For each object physics behaviors can be enabled individually and recorded together with all other forms of animation.

6.3. Camerawork

Cameras can be placed on set as an asset or a special tool. The Viewfinder can be used to frame a shot and create a new camera. Camera views can be previewed from the Wrist Pad. By clicking on a camera, the user can enter the "Through the Lens" mode which effectively puts them in the position of a camera operator. Camera motion is animated to move between keyframes. This emulates dolly tracks, cranes and other staples of camerawork.

Figure 22. View from the "Through the Len" mode.

6.4. Shot Sequencer

One of the key outputs of the main prototype are the shot sequences. The shot sequencer is a desktop tool to assemble recorded shots. After recording a shot, one can select and preview cameras, cut to the chosen camera on the master track, trim/reveal camera clips, and export the final project as video file.

6.5. Collaboration

While many tasks only directly involve one practitioner, one of the main purposes of previs is to communicate ideas amongst the team and manage any issues and risks that arise. The prototype supports *multiple users* working on a scene at the same time, each in their own location. Team members can see avatars of each other and speak to one another. When a team member makes an edit to the scene, other team members see the effect of that edit as the objects are synced within the members.

6.6. Evaluation of the Main Prototype

We aimed to evaluate the main NUI prototype in terms of applicability for the usage in realistic projects

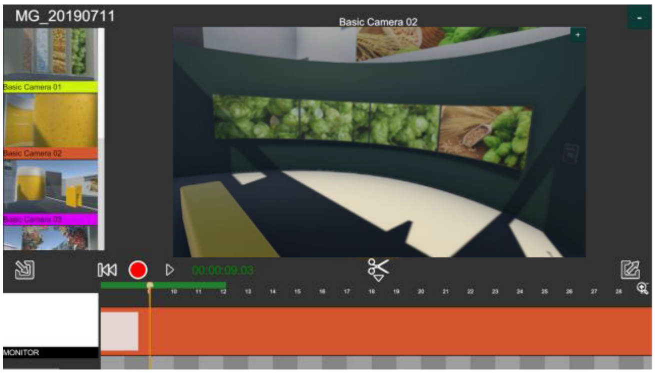

Figure 23. After a shot has been recorded, various editing functions can be accessed to produce the final result.

from the creative industries of film, animation and theater. For that purpose, three project scenarios were defined with professionals from the respective creative domains. For each project, the scope and overall objective were defined based on actual finished or ongoing projects that were developed in the experts' companies. For the realization of these projects, the studios used the NUI prototype and conducted the development process autonomously. Once the projects were completed successfully, we evaluated the procedure with the help of semi-structured interviews. Here, we focused on collecting qualitative data to get an insight into the overall user experience of using NUIs for previs. In total, we presented nine questions to the professionals, for instance asking if they were able to achieve their goals using the prototype or how the workflow compared to their traditional previs procedure. Across all examined application domains, we gathered highly positive feedback. The industry experts involved in the interviews emphasized that they would benefit greatly by using a software based on NUIs in their working routines. As the feedback expressed by the experts differed quite substantially dependent on the respective domain, we will now provide a short overview for each project individually. For detailed information regarding the responses of the interviews, please refer to the corresponding case study [6].

Animation. For the animation project, the studio in charge assembled a team of five people with varying backgrounds: a director, a production manager, a layout artist, a motion-capture (MoCap) performer and a pipeline supervisor. Their objective was to shoot a trailer for an animated movie by using the NUI prototype. In the interviews conducted after project completion, the experts pointed out that the software gave them new ways to express themselves visually while offering an intuitive and natural form of interaction. On the downside, interviewees stated that they were not familiar with the prototype and thus, needed some time to adjust to the new working environment. In contrast to that however, they also rated the overall learnability to be higher compared to any traditional software alternative available on the market. Being asked about the reasoning for this assessment, the experts stated that it normally takes a vast amount of time to become a professional with any standard 3D content creation software. With the NUI software however, they were able to internalize all necessary functions within hours. In addition to learnability, another advantage of NUIs was found in the option to create animations in real time and review the results on the spot instead of separating these two steps. By using the NUI software, artists and directors could communicate more directly exchanging ideas and feedback right there in the animation phase. In contrast to these benefits, the interviewees also mentioned that wearing the head-mounted display for the duration of the project was rather tiring and exhausted them after a couple of hours. Even for a small-scale project, previs is a lengthy process and thus, the professionals had to take additional breaks due to the cumbersome nature of the hardware.

Film. The film team included a lighting artist, a set designer, a stage designer and a camera operator who worked together to create previs for a short commercial video for one of the studio's customers. A large variety of positive feedback was expressed by the experts. First off, in accordance with the animation team, learnability of the software was emphasized to be one of its major advantages. Instead of lengthy training, the NUI software only required some familiarization and was rather easy to use. Being asked about specific use cases where the experts would benefit most from using NUIs, they mentioned the planning phase of placing props and choosing the right equipment for the production. Furthermore, they pointed out that previs in the film domain usually requires large sets being built and moved around. Making use of VR however allowed them to freely try out different settings, camera placements and light setups. On top of that, financial advantages were mentioned as well since time spent on a film set is rather costly and thus, digital previs would reduce costs immensely. As a final note, the interviewees concluded that the database-structure of the prototype would greatly enhance collaboration among team members. This new approach allowed people involved in the production to work from different places around the globe simultaneously. A suggestion for improvement was expressed by the camera operator who suggested to have a physical replica of an actual camera in the virtual scene for better haptic feedback.

Theater. For this project, the theater involved a stage master, a lighting artist, a CTO, a project manager, two stage technicians and an event technology trainee. Their objective was to create a collection of scenes to previsualize an opera performance. Previs in theater productions is usually based on analogue elements such as cardboard props. Therefore, using VR and NUIs to work on a production felt new and exciting from the team's perspective. Although the experts were unfamiliar with this approach to previs, they found the software to be usable, clearly structured, easy to learn and overall a convincing alternative to traditional methods, offering a "great experience". One major feature they found particularly engaging was the ability for be fully immersed in a virtual scene instead of working with real props in the outside world. Just like in the film production team, the experts pointed out financial advantages of using such a software since time on stage is rather expensive. Additionally, they benefited strongly from using VR to plan out the lighting setup for corridors and areas on stage that were difficult to illuminate. As a suggestion for further development, some interviewees expressed the need for more assets to be used in the stage preparation.

7. Discussion

In this work, we developed and evaluated prototypes based on our proposed concept of using natural user interfaces (NUIs) built around a set of interaction techniques and concepts. Overall, The feedback collected in the course of the project evaluations was highly positive and supporting. The professionals from the creative industries stated affirmative appraisal of the prototypes pointing out how they would benefit from it in their daily workflows.

To develop a previs software that supports the need of individual creative personnel, we identified the necessary functionalities needed for a dedicated previs software. The accessibility of a previs software can be improved if the identified tasks are supported. Moreover, in order to have high usability, these tasks need to have intuitive controls and not highly rely on prior technical knowledge.

The individual prototypes provided us with valuable insights regarding our NUI concept. For each prototype, we scientifically evaluated individual interaction concepts and principals. These prototypes showed great potential to support creatives in different stages of the previs process. We used the feedback from these studies and integrated them partially or entirely into the main prototype.

Regarding the evaluation of the main prototype, a major aspect expressed by the experts was the ability to collaborate easily with other project members. Participants stated that animating or sketching in real-time and receiving feedback from colleagues instantly is an advantage in comparison to the traditional 3D software. The collaboration aspect is further improved by having the ability to save projects in a database and access the same scene from anywhere in the world. In the modern industry, it is often the case that team members are located in different places. Having such collaborative features can be extremely practical and helpful. Furthermore, participants found the planning for placement of lights and superstructures highly beneficial for their workflow. As this process is usually cumbersome and expensive when conducted in a physical set or stage, the VR previs tool helps to overcome these challenges and provides artists with an opportunity to try different constellations without worrying about the costs.

All participants were able to get to the desired outcome in a reasonable amount of time, even though they had to complete a rather complex task without having prior knowledge about the prototype and some even without being introduced to VR at all. Obtaining similar results using a professional software without long training sessions is very unlikely. We can therefore argue that our NUI-based prototype has the potential to offer a usable and more intuitive interaction style than the common software after a short familiarization period or tutorial has been completed. NUIs and VR are not just limited to the playful applications they are mostly used in today. Our findings indicate that they are applicable to professional workflows and can provide great advantages to creative work.

Taking these results into account, we will address some of our interaction principles that were presented in the early stages of the project. Within the course of the project evaluations, we could demonstrate that professional users from the creative industries were able to create digital previs in a fast and easy way having generated accurate and precise end results. The artists and creatives were able to directly interact with the software and exhibit their ideas. In our main prototype evaluations, all participants were capable of completing the test cases they were given and thus, generate convincing previs in the course of mere minutes. Additionally, expert users stated the potential of the software and pointed out its ease of use. Participants also addressed the entertaining factor of using the prototype and how they enjoyed the time they spent interacting with the software. Therefore, we conclude that the "Easy to Use", "Rapidness, Accuracy and Precision", "Playful and Fun" and "Creativity Support" interaction principles have been successfully integrated into the prototype. The main prototype also supports a "Multi-User" environment by providing the possibility to collaborate easily with other project members from various places around the world. Multiple users being able to access the same scene from different locations and make changes on a project at the same time was expressed by the experts.

7.1. Limitations

Due to the large amount of features presented in the main prototype, some participants felt overwhelmed to an extent which led to a high mental workload and a medium assessment of usability. However, we should keep in mind that the subjects had to perform a set of complex tasks without knowing the software and without having any prior experience in VR. For the future iterations of the software, we intend to include a tutorial-mode, slowly guiding through each functionality of the prototype. This mode could help to educate untrained personnel which would result in a more efficient, less time-consuming usage of the software overall.

Fatigue after long sessions was mentioned as another point of criticism. Especially wearing an HMD for longer periods of time was perceived to be rather cumbersome. Overcoming this problem poses quite the challenge for software developers since the hardware dictates wearing comfort and thus, how physically demanding the interaction can become. However, one potential solution could be to offer desktop-alternatives to some of the functionalities. This way, users could take breaks from VR in between long work sessions while still being able to work on their product. Certainly, this would take away the naturalness of the interaction for these periods. Nevertheless, it could be a way to prevent physical exhaustion and on top of that, empower users further by providing different options to use the software.

8. Conclusion

In this article, we introduced a NUI concept for previsualization in an attempt to empower artists and practitioners to intuitively visualize their ideas and show their capabilities depending on the task they want to perform in the previs cycle. Initially, requirements for a previs software were collected from experts in the application areas and the core functionalities were defined based on those requirements. In addition, a concept for the interaction with the system based on natural user interfaces was introduced to guide the development of the user interface for the main prototype as well as exploratory research prototypes. Our findings suggest that NUIs can provide an applicable alternative to traditional design tools and only needs a brief familiarization phase instead of lengthy training as professional software demands. For non-technical personnel, this approach might not only offer an alternative, but a more usable and empowering tool than current software solutions. Overall, the application of VR for previs seems to be beneficial for creative personnel with little technical knowledge due to its easy and intuitive use.

For future developments, we plan to include other modalities that might improve the naturalness of the interaction and further utilize the feedback from the experts to refine the software prototype. As stated by the professionals, haptic feedback plays a major role in certain contexts. For this purpose, we intend to identify where and how to include a more tangible interaction technique. Moreover, we plan to investigate the usefulness of speech and gesture recognition for specific tasks.

Acknowledgement. This project has received funding from the European Union's Horizon 2020 research and innovation programme (No 688244). We thank all partners who contributed to this research.

References

[1] Yamaguchi, S., Chen, C., Yoon, D. and Gregoire, D. (2015) Previsualization: How to develop previs in asia? In *SIGGRAPH Asia 2015 Symposium on Education*, SA '15 (New York, NY, USA: ACM): 16:1–16:3. doi:10.1145/2818498.2818515, URL `http://doi.acm.org/10.1145/2818498.2818515`.

[2] Wong, H.H. (2012) Previsualization: Assisting filmmakers in realizing their vision. In *SIGGRAPH Asia 2012 Courses*, SA '12 (New York, NY, USA: ACM): 9:1–9:20. doi:10.1145/2407783.2407792, URL `http://doi.acm.org/10.1145/2407783.2407792`.

[3] Muender, T., Fröhlich, T. and Malaka, R. (2018) Empowering creative people: Virtual reality for previsualization. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18 (New York, NY, USA: ACM): LBW630:1–LBW630:6. doi:10.1145/3170427.3188612, URL `http://doi.acm.org/10.1145/3170427.3188612`.

[4] Nitsche, M. (2008) Experiments in the use of game technology for pre-visualization. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, Future Play '08 (New York, NY, USA: ACM): 160–165. doi:10.1145/1496984.1497011, URL `http://doi.acm.org/10.1145/1496984.1497011`.

[5] Muender, T., Volkmar, G., Wenig, D. and Malaka, R. (2019) Analysis of previsualization tasks for animation, film and theater. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19 (New York, NY, USA: ACM): LBW1121:1–LBW1121:6. doi:10.1145/3290607.3312953, URL `http://doi.acm.org/10.1145/3290607.3312953`.

[6] Volkmar, G., Muender, T., Wenig, D. and Malaka, R. (2020) Evaluation of natural user interfaces in the creative industries. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*: 1–8.

[7] Vetere, F., O'Hara, K., Paay, J., Ploderer, B., Harper, R. and Sellen, A. (2014) Social nui: Social perspectives in natural user interfaces. In *Proceedings of the 2014 Companion Publication on Designing Interactive Systems*, DIS Companion '14 (New York, NY, USA: ACM): 215–218. doi:10.1145/2598784.2598802, URL `http://doi.acm.org/10.1145/2598784.2598802`.

[8] Fu, L.P., Landay, J., Nebeling, M., Xu, Y. and Zhao, C. (2018) Redefining natural user interface. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18 (New York, NY, USA: ACM): SIG19:1–SIG19:3. doi:10.1145/3170427.3190649, URL http://doi.acm.org/10.1145/3170427.3190649.

[9] García-Peñalvo, F.J. and Moreno, L. (2019) Special issue on exploring new natural user experiences. *Universal Access in the Information Society* **18**(1): 1–2. doi:10.1007/s10209-017-0578-0, URL https://doi.org/10.1007/s10209-017-0578-0.

[10] Naumann, A., Hurtienne, J., Israel, J., Mohs, C., Kindsmüller, M., Meyer, H. and Husslein, S. (2007) Intuitive use of user interfaces: Defining a vague concept: 128–136. doi:10.1007/978-3-540-73331-7_14.

[11] Clifford, R.M.S. and Billinghurst, M. (2013) Designing a nui workstation for courier dispatcher command and control task management. In *Proceedings of the 14th Annual ACM SIGCHI-NZ Conference on Computer-Human Interaction*, CHINZ '13 (New York, NY, USA: ACM): 9:1–9:8. doi:10.1145/2542242.2542252, URL http://doi.acm.org/10.1145/2542242.2542252.

[12] Bolt, R.A. (1980) "put-that-there" voice and gesture at the graphics interface. In *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*: 262–270.

[13] Francese, R., Passero, I. and Tortora, G. (2012) Wiimote and kinect: Gestural user interfaces add a natural third dimension to hci. doi:10.1145/2254556.2254580.

[14] Delimarschi, D., Swartzendruber, G. and Kagdi, H. (2014) Enabling integrated development environments with natural user interface interactions. In *Proceedings of the 22Nd International Conference on Program Comprehension*, ICPC 2014 (New York, NY, USA: ACM): 126–129. doi:10.1145/2597008.2597791, URL http://doi.acm.org/10.1145/2597008.2597791.

[15] Jain, J., Lund, A. and Wixon, D. (2011) The future of natural user interfaces. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, 211–214.

[16] Weiyuan Liu (2010) Natural user interface- next mainstream product user interface. In *2010 IEEE 11th International Conference on Computer-Aided Industrial Design Conceptual Design 1*, **1**: 203–205. doi:10.1109/CAIDCD.2010.5681374.

[17] Jacob, R.J., Girouard, A., Hirshfield, L.M., Horn, M.S., Shaer, O., Solovey, E.T. and Zigelbaum, J. (2008) Reality-based interaction: A framework for post-wimp interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08 (New York, NY, USA: ACM): 201–210. doi:10.1145/1357054.1357089, URL http://doi.acm.org/10.1145/1357054.1357089.

[18] Walther-Franks, B. and Malaka, R. (2014) An interaction approach to computer animation. *Entertainment Computing* **5**(4): 271–283.

[19] Lee, J., Chai, J., Reitsma, P.S.A., Hodgins, J.K. and Pollard, N.S. (2002) Interactive control of avatars animated with human motion data. In *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '02 (New York, NY, USA: ACM): 491–500. doi:10.1145/566570.566607, URL http://doi.acm.org/10.1145/566570.566607.

[20] Chai, J. and Hodgins, J.K. (2005) Performance animation from low-dimensional control signals. In *ACM SIGGRAPH 2005 Papers*, 686–696.

[21] Walther-Franks, B., Biermann, F., Steenbergen, N. and Malaka, R. (2012) The animation loop station: Near real-time animation production. In *Proceedings of the 11th International Conference on Entertainment Computing*, ICEC'12 (Berlin, Heidelberg: Springer-Verlag): 469–472. doi:10.1007/978-3-642-33542-6_55, URL http://dx.doi.org/10.1007/978-3-642-33542-6_55.

[22] Walther-Franks, B., Herrlich, M., Karrer, T., Wittenhagen, M., Schröder-Kroll, R., Malaka, R. and Borchers, J. (2012) Dragimation: Direct manipulation keyframe timing for performance-based animation. In *Proceedings of Graphics Interface 2012*, GI '12 (Toronto, Ont., Canada, Canada: Canadian Information Processing Society): 101–108. URL http://dl.acm.org/citation.cfm?id=2305276.2305294.

[23] Igarashi, T., Moscovich, T. and Hughes, J.F. (2005) Spatial keyframing for performance-driven animation. In *Proceedings of the 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '05 (New York, NY, USA: ACM): 107–115. doi:10.1145/1073368.1073383, URL http://doi.acm.org/10.1145/1073368.1073383.

[24] Liu, C.K. and Zordan, V.B. (2011) Natural user interface for physics-based character animation. In *Proceedings of the 4th International Conference on Motion in Games*, MIG'11 (Berlin, Heidelberg: Springer-Verlag): 1–14. doi:10.1007/978-3-642-25090-3_1, URL http://dx.doi.org/10.1007/978-3-642-25090-3_1.

[25] Shum, H. and Ho, E.S. (2012) Real-time physical modelling of character movements with microsoft kinect. In *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology*, VRST '12 (New York, NY, USA: ACM): 17–24. doi:10.1145/2407336.2407340, URL http://doi.acm.org/10.1145/2407336.2407340.

[26] Herrlich, M., Braun, A. and Malaka, R. (2012) Towards bimanual control for virtual sculpting. In Reiterer, H. and Deussen, O. [eds.] *Mensch & Computer 2012: interaktiv informiert – allgegenwärtig und allumfassend!?* (München: Oldenbourg Verlag): 173–182.

[27] Galyean, T.A. and Hughes, J.F. (1991) Sculpting: An interactive volumetric modeling technique. In *Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '91 (New York, NY, USA: ACM): 267–274. doi:10.1145/122718.122747, URL http://doi.acm.org/10.1145/122718.122747.

[28] Chen, H. and Sun, H. (2002) Real-time haptic sculpting in virtual volume space. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '02 (New York, NY, USA: ACM): 81–88. doi:10.1145/585740.585755, URL http://doi.acm.org/10.1145/585740.585755.

[29] Galoppo, N., Otaduy, M.A., Tekin, S., Gross, M. and Lin, M.C. (2007) Soft articulated characters with fast contact handling. *Computer Graphics Forum* **26**(3): 243–253. doi:10.1111/j.1467-8659.2007.01046.x, URL

https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8659.2007.01046.x.

[30] Wesson, B. and Wilkinson, B. (2013) Evaluating organic 3d sculpting using natural user interfaces with the kinect. In *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*, OzCHI '13 (New York, NY, USA: ACM): 163–166. doi:10.1145/2541016.2541084, URL http://doi.acm.org/10.1145/2541016.2541084.

[31] "Vogel, D., Lubos, P. and Steinicke, F. ("2018") "animationvr - interactive controller-based animating in virtual reality". In *"Proceedings of the 1st Workshop on Animation in Virtual and Augmented Environments"* ("IEEE"). URL "http://basilic.informatik.uni-hamburg.de/Publications/2018/VLS18b".

[32] Okun, J., Zwerman, S. and Society, V.E. (2010) *The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures*, Media Technology (Focal Press/Elsevier). URL https://books.google.de/books?id=G_08wSzR4EIC.

[33] Chemuturi, M. (2012) *Requirements Engineering and Management for Software Development Projects* (Springer Publishing Company, Incorporated).

[34] Kwon, B.c., Javed, W., Elmqvist, N. and Yi, J.S. (2011) Direct manipulation through surrogate objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11 (New York, NY, USA: ACM): 627–636. doi:10.1145/1978942.1979033, URL http://doi.acm.org/10.1145/1978942.1979033.

[35] Raskin, J. (2000) *The Humane Interface: New Directions for Designing Interactive Systems* (New York, NY, USA: ACM Press/Addison-Wesley Publishing Co.).

[36] Bérard, F., Ip, J., Benovoy, M., El-Shimy, D., Blum, J.R. and Cooperstock, J.R. (2009) Did "minority report" get it wrong? superiority of the mouse over 3d input devices in a 3d placement task. In *INTERACT*.

[37] Ryan, R.M. and Deci, E.L. (2000) Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* **55**(1): 68.

[38] Deterding, S., Dixon, D., Khaled, R. and Nacke, L. (2011) From game design elements to gamefulness: Defining gamification. **11**: 9–15. doi:10.1145/2181037.2181040.

[39] Norman, D. (2013) *The Design of Everyday Things: Revised and Expanded Edition* (Constellation).

[40] Muender, T., Reinschluessel, A.V., Drewes, S., Wenig, D., Döring, T. and Malaka, R. (2019) Does it feel real?: Using tangibles with different fidelities to build and explore scenes in virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19 (New York, NY, USA: ACM): 673:1–673:12. doi:10.1145/3290605.3300903, URL http://doi.acm.org/10.1145/3290605.3300903.

[41] Fröhlich, T. (2020) Natural and playful interaction for 3d digital content creation .

[42] Fröhlich, T., von Oehsen, J. and Malaka, R. (2017) Productivity & play: A first-person shooter for fast and easy scene design. In *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '17 Extended Abstracts (New York, NY, USA: Association for Computing Machinery): 353–359. doi:10.1145/3130859.3131319, URL https://doi.org/10.1145/3130859.3131319.

[43] Fröhlich, T., Alexandrovsky, D., Stabbert, T., Döring, T. and Malaka, R. (2018) Vrbox: A virtual reality augmented sandbox for immersive playfulness, creativity and exploration. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '18 (New York, NY, USA: ACM): 153–162. doi:10.1145/3242671.3242697, URL http://doi.acm.org/10.1145/3242671.3242697.

[44] Bonfert, M., Porzel, R. and Malaka, R. (2019) Get a grip! introducing variable grip for controller-based vr systems. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (IEEE): 604–612.