EUROGRAPHICS 9th Workshop on Cinematography and Editing (2020)
M. Christie, H.-Y. Wu, T.-Y. Li and V. Gandhi (Editors)

Designing an Adaptive Assisting Interface for Learning Virtual Filmmaking

Qiu-Jie Wu[1], Chih-Hsuan Kuo[1], Hui-Yin Wu[2] and Tsai-Yen Li[1]

[1]National Chengchi University, Taiwan
[2]Université Côte d'Azur, Inria, France

Abstract

In this paper, we present an adaptive assisting interface for learning virtual filmmaking. The design of the system is based on the scaffolding theory, to provide timely guidance to the user in the form of visual and audio messages that are adapted to each person's skill level and performance. The system was developed on an existing virtual filmmaking setup. We conducted a study with 24 participants, who were asked to operate the film set with or without our adaptive assisting interface. Results suggest that our system can provide users with a better learning experience and positive knowledge harvest.

CCS Concepts

• ***Human-centered computing*** → *Interaction design process and methods; User studies;* • ***Computing methodologies*** → *Virtual reality;* • ***Applied computing*** → *Media arts;*

1. Introduction

Film cinematography has the function of eliciting emotions from the audience, which influences not only how scripts are written, but also how the director chooses to direct the movie. The techniques applied by the cinematographer can lead to a wide variety of interpretations of the same story. All of these show the role of cinematography in telling the story of the film. It is important for the users studying film production to be able to experiment with different ways of filming and observe the effects on the audience.

To improve one's skills, we believe that practice makes perfect. However, typically, building a film set with real props and actors is costly and not readily available to most novice filmmakers. Recently, virtual reality has been deployed to develop a virtual film set [LLGC18] that allows users to take on the role of the cinematographer, director, or editor, thus allowing trainees and students to experience the process of filmmaking at a lower cost, and with minimal professional equipment and personnel. Unfortunately, virtual reality technologies introduce a high technical barrier for those with little to no experience in 3D virtual environments. In this work, we design an innovative adapting assisting interface to help the user learn to operate a virtual film set. Our primary goal is to assist the users in familiarizing themselves with the virtual set without the need for in-person mentoring. A secondary goal is to offer non-expert users the opportunities to make a head start on cinematography.

To this end, we refer to the scaffolding theory, introduced in the late 1950s by Jerome Bruner. The term "scaffolding" represents the helpful interactions that enable the students to do something beyond his or her independent efforts [MS08]. Based on this concept, we designed an adaptive assisting learning interface that builds on an existing virtual film set, which tutors the users through visual and audio messages. The user is assigned an initial skill level at the beginning of the tutorial. As they advance through the hands-on tutorial, the system automatically adjusts the level of help according to the user's performance. We carried out a user study to evaluate the assisting interface and scaffolding mechanism that we have designed, the results of which are reported in this paper.

The paper is organized as follows. First, the state of the art on the scaffolding theory and virtual camera control is presented. We then show how the adaptive assistant is constructed, and our system architecture. This is followed by the description and results of a study in which we invited users to evaluate our system. Finally, we summarise the findings and limitations and offer perspectives on future work.

2. Related Work

In this section, we provide an overview of research on the scaffolding theory and the camera control methods for 3D environments that concern the design and operation of the virtual film set.

2.1. Intelligent tutoring systems and virtual training

For many years, virtual reality (VR) has been expected to bring on the next evolution in learning technologies, introducing several advantages, such as low cost, high comfort, and unlimited distance [PIB*16, RBB*00, You98]. In a recent work on adaptive technologies in virtual reality training [VGD16], data from trainee assessment was used as an input to an autonomous system for customized training and automated difficulty level adjustment to meet individual needs.

2.2. Scaffolding Theory

Scaffolding theory was first introduced in the late 1950s by Jerome Bruner, a cognitive psychologist. Scaffolding represents the helpful interactions that enable students to do something beyond his or her independent efforts [MS08]. Instructional scaffolding is the support given to a student by an instructor throughout the learning process. This support is specifically tailored to each student such that the "scaffolds" are incrementally removed when the students acquire the knowledge or skills under the learning target. This instructional approach allows students to experience student-centered learning, which tends to facilitate more efficient learning than teacher-centered learning.

In the learning process, the educators may question a student's approach to a difficult problem and provide constructive feedback. The type and amount of support needed are dependent on the needs of the students during the time of instruction [VL14]. Scaffolds are developed to assist students with a difficult task [SB02]. The key is that the assistance needs to be planned.

Scaffolding is also used to support problem-based learning (PBL), an instructional approach in which students are presented with an ill-structured problem [Ali19]. After being presented with the problem, defining it, and generating learning issues, students proceed to address their learning issues, and then develop a potential solution, backed with evidence. There is some evidence that PBL can be effective when students work individually, leading to strong learning outcomes. Many educators incorporate PBL in their classrooms to engage students and help them become better problem solvers. However, educators must first identify the content that needs scaffolding support, the appropriate time to implement the support, and determine when the scaffold can be removed [Laj05].

The design of our interface was also inspired by game-based learning, which can help learners acquire real-world knowledge through a game-like environment. A study by Barzilai and Blau [BB14] suggests that presenting scaffolds may have "problematized" learners' understandings of the game by connecting them to disciplinary knowledge. Implications for the design of scaffolds for game-based learning are discussed. Zook et al. [DZO*13] also demonstrate how a rule-based intelligent system can reduce the frequency of errors that novices make by providing information about rule violations without prescribing solutions, and discuss the role of error reduction in creativity support tools.

2.3. Camera Control Methods

In computer graphics and animations, autonomous camera control has been a challenge to find from infinite possibilities to film a set, solutions that fulfill geometrical, storytelling, and navigation

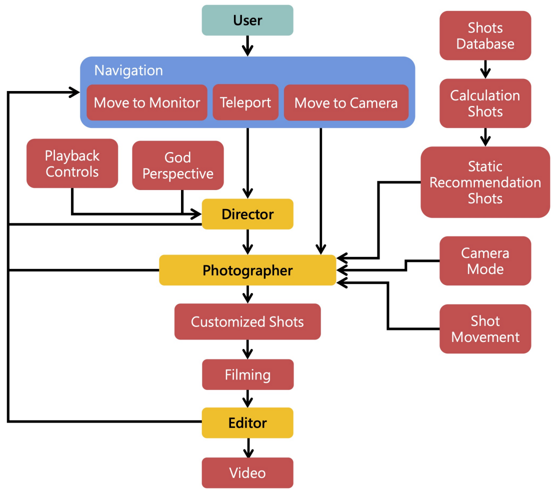

Figure 1: *The architecture of our virtual film set is composed of three main parts: (1) a navigation component to move around cameras and the monitor, (2) shot management components (right of the figure) including a database, recommendation, and editing functions, (3) the workspace allowing users to take on various filmmkaing roles, and export the output as a video.*

constraints. Christie et al. [CO09] have summarized comprehensive solution techniques from different approaches. While photography and cinematography practices have been documented for the operation of real-world cameras, the cinematographers and photographers have little to no control over the staging of day-to-day content. Whereas in 3D environments, cameras can be intuitively placed at given coordinates and rotated. Solutions to generate camera paths given constraints have also been proposed [GCLR15].

For virtual cinematography [HCS96], Jhala and Young created the Darshak system that generates shot sequences to fulfill goals of showing specific actions and events in a 3D environment [JY11]. O'Neill et al. [ORN09] proposed a system of blockings that would choreograph the relative positions of actors and cameras in an environment. In these systems, context – the placement of characters and cameras in accordance with the story and the environment – is crucial, making camera control in 3D environments a complex problem.

The adoption of cinematic idioms for virtual camera control was first demonstrated through rule- and constraint-based systems [CAH*96, BGL98], and through the representation of state transitions between various shot types [HCS96]. Ronfard et al. then proposed the Prose Storyboard language (PSL) [RGB13] as a formal representation for describing shots visually. PSL illustrates the spatial and temporal structure of movie frames with arbitrarily complex shots. For instance, each shot in a movie consists of a variety of sizes and angles, and the visual composition of every frame is defined as relative position and orientation by applying the prose storyboard language. The Director's Lens [LCRB11] was the first to introduce a mixed-initiative creative interface (MICI) with viewpoint suggestions and a tangible camera device to control static and dynamic cameras. Further along, optimization approaches [GRLC15] and film editing patterns [WPRC18] were proposed to ensure spatial and temporal continuity, and to apply common film idioms over long sequences of shots respectively.

3. System Framework

3.1. Virtual Film Set

In this section, we describe the design of our system. To understand our approach, we first briefly introduce the virtual film set that has been developed in previous work [LLGC18]. We then describe the design of the adaptive assisting interface for learning that is built on the virtual film set.

The architecture of our virtual film set, depicted in Figure 1, is composed of three main parts: (1) a navigation component to move around cameras and the monitor, (2) shot management components including a database, recommendation and editing functions, and (3) the workspace allowing users to take on various filmmaking roles, and export the output to a video. The system was developed using Unity 3D and runs on HTC VIVE. While operating the virtual film set, a user simply puts on the headset and interacts with the set using the two hand controllers.

In this virtual film set, a user can take on the role of either the director, photographer, or editor. Each of these represents important roles in the filmmaking process. As a director, the user can decide the number of cameras that should be placed by observing the whole film set from the god's perspective and the storyboard; as a photographer, the user can determine where to place the cameras to have the desirable framing; as an editor, the user is expected to arrange the story by composing a sequence of shots from the various camera streams.

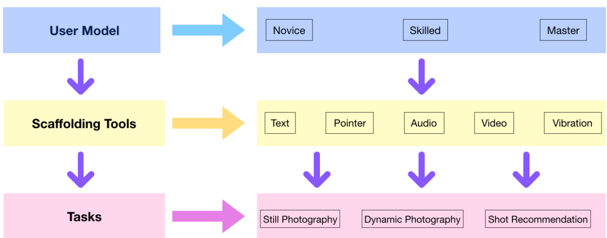

Figure 2: *Our assisting learning interface has three modules: the user model, scaffolding tools, and tasks*

In this work, the capabilities of the photographer are enhanced through the design of a recommendation component for static camera shots. The recommendation is based on a shot database built from the experience of filmmaking experts. We invited eight professional photographers to shoot a sequence on the reconstructed animation from a clip in *Back to the Future*, and analyzed their shot compositions on the keyframes. Finally, we added the results to the shot database to improve the quality of the shot recommendation function provided by our system.

3.2. Assisting Learning Interface

Although the goal of design the proposed virtual film set system is to help users learn or experience filmmaking through virtual reality, the user has to learn how to use the system first. Therefore, we have designed a user interface that can assist users in learning such a system. The assisting interface has three main modules: the user model, scaffolding tools, and tasks as shown in Figure 2. The system divides users into three levels: novice, skilled, and master, based on the user's self-evaluation and their understanding of the virtual film set.

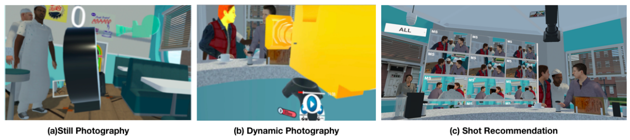

Figure 3: *Three tasks have been designed to illustrate the assisting learning interface as well as to help the users learn the use of virtual film set: still photography, dynamic photography, and the use of shot recommendation*

According to instructional scaffolding, users of different skill levels should receive different levels of hints (scaffolds). In our current system, we have provided five prompt methods as tools, as shown in Figure 4, to assist users in learning how to use the system :

- Text hint: tells the user the specific operation mode through text
- Pointer hint: lets the user know the position of the button for the execution of the next step by showing a pointer prompt
- Audio hint: helps users by giving audio information
- Video hint: demonstrates the correct operation method and sequence through a video
- Vibration hint: tells the user when they have completed the task by giving the user feedback through the vibration of the HTC VIVE's controllers

Different instructional tools are adopted in different tasks according to the context and difficulty of the task. We have designed several different types of tasks to illustrate how assisting learning interface can be designed to fit the need of the tasks. In our current system, three main tasks are available to the users: still photography, dynamic photography, and the use of shot recommendation as shown in Figure 3. Each task may be further divided into several subtasks (Table 1), which are smaller learning units that are easier for the users to master through practice.

When the user starts a task, the assisting learning interface will start timing. In the whole learning process, we have three important time points (denoted by t_1, t_2, t_3 in the following text).

If the time reaches t_1, and the user has not started to do anything, the assisting learning interface will give the user a text hint, which will introduce the user to the task in detail again and instructions on how to do it.

After t_2 seconds, if the user still has not completed the task, the assisting learning interface will give the user a pointer hint to show

Table 1: *Hints for each subtask*

Tools / Tasks	Basic Control			Still Photography				Dynamic Photography	
	Movement	Teleport	Grab	Add Camera	Delete Camera	Move Camera	Shot Recommendation	Play Media	Track Actor
Text	✓	✓	✓	✓	✓	✓	✓	✓	✓
Step					✓	✓	✓	✓	✓
Audio	✓	✓	✓	✓	✓	✓	✓	✓	✓
Video	✓	✓	✓			✓	✓		✓
Vibration	✓					✓			✓

which button to press next, the position of the button, and the function of the button.

If the user still has not completed the task by t_3, the assisting learning interface will give the user an audio hint. Through a voice, the assisting learning interface will guide the user to complete the task as a teacher may do besides the student.

In our design, some hints appear in all tasks. According to the learning context and difficulty of the task, the assisting function is either enabled or optional. Currently, three hint functions (text, pointer, audio) are the basic ones for all tasks while the other hints are used to provide further scaffolds when needed. For example, the video hint is provided when the operations used in completing a task are more complex and cannot be described clearly with text. In this case, showing a demonstration video may help the user understand the specific operations that need to be carried out.

Besides, when the system wants to provide feedback to the user upon completion of an operation or task, the vibration function can be used. Further specific information about the use of hints in different tasks are given in Table 1 .

When a user enters our system to learn a specific task in the virtual film set, the system will give hints (scaffolds) to the user according to the task as well as the proficiency of the user. The hints appear progressively. For example, usually, a text hint appears first after t_1 seconds of delay. When the task consists of several operations, the pointer hint will then appear t_2 seconds after the text hint. If the user still cannot accomplish the task, the audio hint will follow after t_3 seconds of delay. For different tasks and different levels of users, these delay parameters will be given different values. For example, for users of the novice level, t_1 is equal to 0 seconds, which means that a text prompt will be given immediately when they start the task. For skilled and master users, t_1 is set to 4 and 6 seconds, respectively. In other words, the hints will be prompted at a later time for more proficient users as they may not need the scaffold to finish the task. Similarly, the ranges of t_2 and t_3 are 2 to 6 seconds and 4 to 8 seconds, respectively. These parameters are set by the designer according to their empirical experience and may vary slightly for different tasks.

3.3. Adaptive Mechanism

According to the principle of instructional scaffolding, the assisting mechanism provided to the users should be customized and adjusted dynamically as their proficiency improves. In our system, we have designed an adaptive mechanism to adjust the level of our assisting learning interface adaptively.

In the adaptive mechanism, the user can provide active feedback to the frequency of the hints such that the level of user proficiency can be adjusted immediately. In the assisting learning interface, we have designed a 'help' button and an 'annoying' button for the user to provide their feedback on the frequency of the hints. When the user encounters difficulties in using the system, he can press the 'help' button, and the system will give him a hint. When the user feels that there are too many prompts and feels 'irritated', he can press the 'annoying' button to reduce the frequency of prompts.

After a task is completed, the level of user proficiency will be adjusted if needed according to two criteria: (1) the time that the user took to complete the last task, and (2) the number of times the user provides feedback by pressing the 'help' and the 'annoying' buttons.

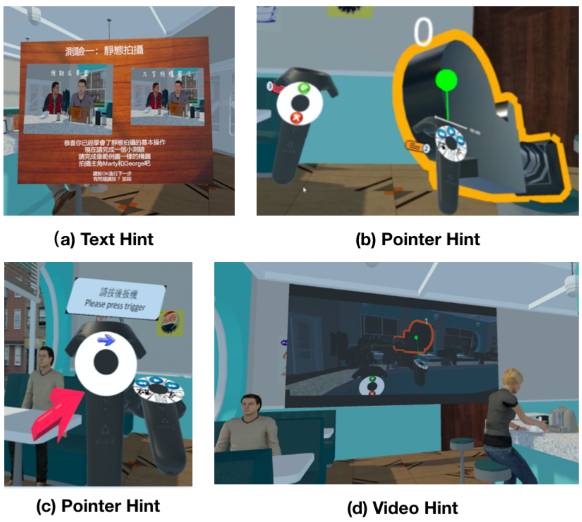

Figure 4: *Our system provides five kinds of hints to the user: text, pointer, audio, video, and vibration. This figure shows examples of the text, video, and pointer hints.*

We also studied the quality of the accomplished task of users and found that most of the user tasks are completed very well, so we decided to use their completion time as a more important basis for judging their proficiency. The system records the time T_c that a user takes to complete a task and uses it to determine if the level needs to be adjusted. The decision is based on the standard completion times T_s for each level of users obtained according to experiments conducted in a pilot study. If T_c is closer to the T_s of a level than others, we will adjust the user's proficiency to this level accordingly. Different tasks for users of different levels of proficiency will have different values of T_s as the difficulty of a task can vary greatly, which affects the completion time.

In addition to the user's completion time, the times that the users press the 'help' and the 'annoying' buttons will also affect the adjustment of level. If the number of times that the 'annoying' button is pressed is over a threshold, the proficiency level will be raised by one level. On the other hand, if the help button is pressed too often, the level will be lowered. Among these two types of criteria, the active feedback provided by the user takes precedence over the completion time since we think it is more subjective and should take priority.

4. Experiments

To validate our system design and receive feedback on the effectiveness of our system, we designed an experiment in which recruited users were asked to carry out several tasks on the virtual film set under different conditions of our assisting learning interface. In this experiment, twenty-four participants (16 females and 8 males) with diverse backgrounds were recruited online. Their ages ranged from 18 to 32, and 9 of them had never experienced virtual reality before. We will detail the procedure and task design, and the experimental results in this section.

4.1. Procedure

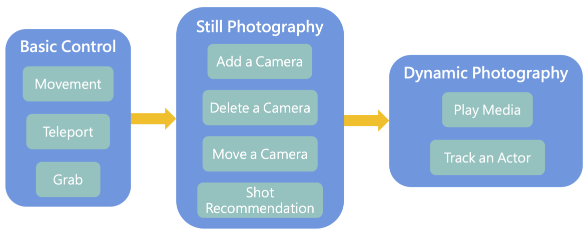

Figure 5: *Here we depict the main process of our experiment, which is composed of three tasks: (1) the basic controls task such as moving around and grabbing objects on the set, (2) still photography task of manipulating a static cameras, and (3) dynamic photography task of playing animations and tracking moving actors.*

First of all, a pre-experiment survey of the experiences of the participants in cinematography and using VR related systems was carried out. At the beginning of the experiment, the participant would select a level that best described how familiar he or she was with the virtual film set. During the experiment, the participant was asked to finish three main tasks which would teach them how to film a sequence in the virtual film set. The main tasks were composed of knowledge about the basic controls, still photography, and dynamic photography, and each of which may consist of multiple sub-tasks.

Figure 6: *The relationship between prior-knowledge and experiment time shows that prior-knowledge did not seem to affect the time it took to complete tasks.*

As shown in Figure 5, the first task is for the participant to practice the basic controls in the virtual film set, including moving around, selecting items, and grabbing objects by using VIVE's controllers in the virtual environment. The next task is to learn how to use the still photography functions, such as adding, deleting, and moving a camera. Also, the participant is asked to try out the new shot recommendation function that is first introduced in this work. The last task is to instruct the participants on how to use dynamic photography functions, including playing the actors' animation and tracking a moving actor with a camera in the virtual film set.

We adopted a between-subject design in our experiment and divided the participants into three groups. The first group was the control group that was only offered the step-by-step tutorial outside the virtual environment, which means that the participant would not see any instructions inside the virtual film set. The second group was the first experimental group that was provided assisting learning interface with a fixed setting. This means that the users in this group did not have access to the adaptive functions and their level of proficiency would remain at the initial setting and would not change throughout the experiment. The third group was the second experimental group that was provided with both the assisting learning interface as well as the adaptive mechanism described in the previous section.

4.2. Experimental Results

Below we analyze the results of the experiment on our system concerning the user's prior knowledge, the performance for each group, and finally the user feedback via a post-experiment survey.

4.2.1. Prior knowledge

In Table 2, we show the statistics of the three groups of participants on their prior knowledge (K), perceived difficulty levels (D), and the time they used in the experiment (T). We used questionnaires

Table 2: *Comparison of prior knowledge, perceived difficulty, and experimental time (AVG (STD))*

	Prior Knowledge Level (K)	Perceived Difficulty Level (D)	Experimental Time (min)(T)
1st Group	5.13 (1.36)	16.88 (2.03)	11.50 (3.38)
2nd Group	4.00 (1.6)	14.88 (2.36)	13.63 (1.85)
3rd Group	3.63 (1.92)	14.87 (2.85)	12.88 (3.4)
Average	**4.25**	**15.54**	**12.67**

to acquire prior knowledge and perceived difficulty. The questions about the prior knowledge include the level of understanding of VR, and basic knowledge of camera shooting.

We use a 10-point scale for the participants to rate, with 1 meaning the user has no prior knowledge and 9 meaning expert. The perceived difficulty of the tasks is based on the subjective evaluation of the difficulty of completing all three tasks on a scale of 1 to 20 points from easiest to most difficult. The null t-test is applied to the prior knowledge of the three groups of users and the result shows that there is no significant difference among them despite that their averages differ. We have also found that, for the three groups, the correlation coefficients for K vs. D are all below 0.4, which means that there is no strong relationship between the prior knowledge and perceived task difficulty.

The prior knowledge (K) and the experimental time (T) for each participant are plotted in Figure 6. As one can observe, there is no direct relationship between prior knowledge and experimental time. The correlation coefficients between K and T for all three groups are below 0.4. Therefore, we believe that users' evaluation of tasks was not affected by their prior knowledge and should be credible.

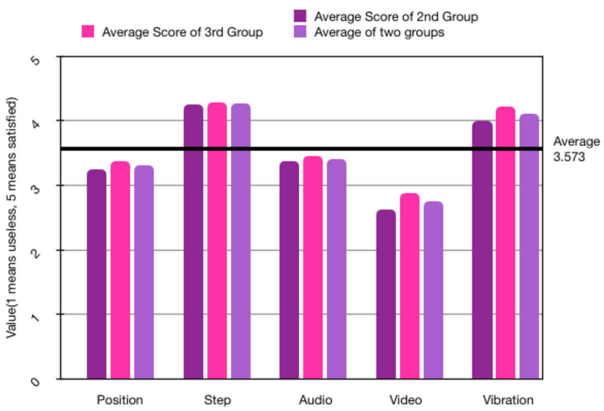

Figure 7: *Evaluation on different types of hints in the assisting learning interface*

4.2.2. Learning Assisting Function

As shown in Table 1, different tasks may adopt different types of hints. To see the effects of different hints in the assisting learning interface, we have asked the participants in the experimental groups (groups 2 and 3) to provide feedback on these hints with a questionnaire after the experiment is finished.

Table 3: *The user reported scores for the five types of hints provided by our assisting learning interface show that the pointer and vibration hints were generally most well received by both groups.*

	Average score of the 2nd Group	Average score of the 3rd Group	**Average score of both groups**
Text	3.25	3.38	**3.31**
Pointer	4.25	4.28	**4.26**
Audio	3.38	3.45	**3.41**
Video	2.63	2.88	**2.75**
Vibration	4.00	4.23	**4.11**
Average	**3.56**	**3.71**	

Table 3 shows the second and third groups of users' evaluations of the five types of hints provided by the assisting learning interface. The evaluation was composed of a list of questions ranked on the 5-point Likert scale with 1-5 meaning mostly disagree to mostly agree. According to Table 3, the average score of the five types of hints is 3.5, which indicates that the assisting learning interface is effective in general.

Among these types of hints, the participants seem to be most satisfied with the pointer hint and the vibration hint, and they think that these are effective for learning virtual film sets. The average score for the video hint is 3 points, which is lower than expected, which means that video hint is not very effective. According to the user's feedback after filling the questionnaire, the sound of the video seems to be too low, and the speed is too fast to clearly understand the content. This problem will be corrected in our future work.

We also found that the pointer hint received the highest score. The participants considered that this hint helped them effectively, especially in complex multi-step operations.

4.2.3. User Feedback

We asked the participants in the two experimental groups to rank their overall experience with the system on a Likert scale of 1 to 5 with 1 for most disagree and 5 for most agree on three statements: (1) With the assisting learning interface, I can easily understand the operations of the virtual film set, (2) I would like to learn the virtual film set through this assisting learning interface, (3) I think

Table 4: *User Feedback on 4 questions in the post-experiment survey between the 2nd and 3rd user groups show that users found the assistive functions useful in familiarizing with the virtual film set.*

		With the assisting learning interface, I can more easily understand the operations of the virtual film set (1=disagree, 5=agree)	I would like to learn the virtual film set through this assisting learning interface again (1=disagree, 5=agree)	I think the virtual tutoring functions can appear when I need them (1=disagree, 5=agree)	I consider the frequency that the virtual tutoring functions appear to be (1=too infrequent, 5=too often)
2nd Group	**Average**	4.00	4.25	4.00	3.50
	Std. Dev.	(0.93)	(0.71)	(0.93)	(0.53)
3rd Group	**Average**	4.25	4.25	4.13	3.00
	Std. Dev.	(0.71)	(0.46)	(0.64)	(0.53)

the virtual tutoring functions can appear when I need them. Table 4 shows the average and standard deviation of the user responses. It is clear from the results that the third group responded more positively towards the statements. Also, the standard deviation of each question is lower than that of the second group. We think this indicates that the adaptive mechanism is an effective way of learning the operations of the virtual film set.

An additional question about the participants' opinion on the frequency that the hints appear was also asked at the end of the survey. The scale of 1 means too infrequent, and 5 means too often. As shown in Table 4, the average score for the second group is 3.5 points while it is 3.0 for the third group. The third group of participants found the frequency of hints to be most appropriate which shows that the adaptive mechanism is indeed effective.

5. Discussion

The main limitation of this work, which is also an ongoing challenge in the area of assisted creativity and smart learning tools, is to correctly identify the user's skill level and barriers, and provide guidance and tips that are personalized to their learning goals. Though similar systems have been proposed such as for computer programming [PZD*18], there is very little work on smart tutoring systems for domains such as filmmaking that have both higher technical barriers and also artistic freedom, and evaluation of such systems is a great challenge.

Another aspect worth noting is the type of hints that are suitable for such an assisting interface. In our experiment, we have observed that pointer hints which highlight the next step in the task, and vibration hints for task completion have been best-received by the users, whereas video hints are very dependent on the quality of the content in the videos. This seems to indicate that the most useful hints are not necessarily the ones containing the most information, but rather those that give the user small indications and encouragements while leaving them with more time to develop their hands-on solutions to the task.

6. Conclusion

We have proposed an adaptive assisting interface for learning virtual filmmaking based on the scaffolding theory. Our system guides the user in operating a virtual film set, by providing text and vocal messages that are adapted to each person's skill level and performance. Through a user evaluation with 24 participants, we see that our system has strong potentials for training on virtual film sets, and can provide users with a better learning experience and fruitful learning outcome.

We hope these first steps explored in this work can open new avenues of research in using virtual reality for smart filmmaking environments, that would also have strong impacts towards assistive creativity in other professional and artistic domains.

7. Acknowledgement

This research was supported by the Ministry of Science and Technology, Taiwan, R.O.C., under contract no. MOST107-2221-E-004-008- and MOST108-2221-E-004-007-MY3.

References

[Ali19] ALI S. S.: Problem based learning: A student-centered approach. *English language teaching 12*, 5 (2019), 73–78. 2

[BB14] BARZILAI S., BLAU I.: Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education 70* (01 2014), 65–79. doi:10.1016/j.compedu.2013.08.003. 2

[BGL98] BARES W. H., GRÉGOIRE J. P., LESTER J. C.: Realtime constraint-based cinematography for complex interactive 3d worlds. In *AAAI/IAAI* (1998), pp. 1101–1106. 2

[CAH*96] CHRISTIANSON D. B., ANDERSON S. E., HE L.-W., SALESIN D. H., WELD D. S., COHEN M. F.: Declarative camera control for automatic cinematography. In *AAAI/IAAI, Vol. 1* (1996), pp. 148–155. 2

[CO09] CHRISTIE M., OLIVIER P.: Camera control in computer graphics: Models, techniques and applications. In *ACM SIGGRAPH ASIA 2009 Courses* (New York, NY, USA, 2009), SIGGRAPH ASIA '09, Association for Computing Machinery. URL: https://doi.org/10.1145/1665817.1665820, doi:10.1145/1665817.1665820. 2

[DZO*13] DAVIS N., ZOOK A., O'NEILL B., HEADRICK B., RIEDL M., GROSZ A., NITSCHE M.: Creativity support for novice digital filmmaking. In *Conference on Human Factors in Computing Systems - Proceedings* (04 2013), pp. 651–660. doi:10.1145/2470654.2470747. 2

[GCLR15] GALVANE Q., CHRISTIE M., LINO C., RONFARD R.: Camera-on-rails: automated computation of constrained camera paths. In *Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games* (2015), pp. 151–157. 2

[GRLC15] GALVANE Q., RONFARD R., LINO C., CHRISTIE M.: Continuity Editing for 3d Animation. In *AAAI Conference on Artificial Intelligence* (Austin, Texas, United States, Jan. 2015), AAAI Press. URL: https://hal.inria.fr/hal-01088561. 3

[HCS96] HE L.-W., COHEN M., SALESIN D.: The virtual cinematographer: A paradigm for automatic real-time camera control and directing. In *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics* (01 1996), pp. 217–224. doi:10.1145/237170.237259. 2

[JY11] JHALA A., YOUNG R. M.: *Intelligent Machinima Generation for Visual Storytelling*. Springer New York, New York, NY, 2011, pp. 151–170. URL: https://doi.org/10.1007/978-1-4419-8188-2_7, doi:10.1007/978-1-4419-8188-2_7. 2

[Laj05] LAJOIE S. P.: Extending the scaffolding metaphor. *Instructional Science 33*, 5-6 (2005), 541–557. 2

[LCRB11] LINO C., CHRISTIE M., RANON R., BARES W.: The director's lens: An intelligent assistant for virtual cinematography. In *MM'11 - Proceedings of the 2011 ACM Multimedia Conference and Co-Located Workshops* (11 2011), pp. 323–332. doi:10.1145/2072298.2072341. 3

[LLGC18] LIN I.-S., LI T.-Y., GALVANE Q., CHRISTIE M.: Design and evaluation of multiple role-playing in a virtual film set. In *Proceedings of the 16th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry* (New York, NY, USA, 2018), VRCAI '18, Association for Computing Machinery. URL: https://doi.org/10.1145/3284398.3284424, doi:10.1145/3284398.3284424. 1, 3

[MS08] MADEIRA C.-A., SLOTTA J. D.: The cambridge handbook of the learning sciences (r. keith sawyer, ed., 2005). *Curriculum Inquiry 38*, 4 (2008), 473–476. URL: https://doi.org/10.1111/j.1467-873X.2008.00425.x, arXiv:https://doi.org/10.1111/j.1467-873X.2008.00425.x, doi:10.1111/j.1467-873X.2008.00425.x. 1, 2

[ORN09] O'NEILL B., RIEDL M. O., NITSCHE M.: Towards intelligent authoring tools for machinima creation. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, 2009), CHI EA '09, Association for Computing Machinery, p. 4639–4644. URL: https://doi.org/10.1145/1520340.1520713, doi:10.1145/1520340.1520713. 2

[PIB*16] PARMAR D., ISAAC J., BABU S. V., D'SOUZA N., LEONARD A. E., JÖRG S., GUNDERSEN K., DAILY S. B.: Programming moves: Design and evaluation of applying embodied interaction in virtual environments to enhance computational thinking in middle school students. In *2016 IEEE Virtual Reality (VR)* (March 2016), pp. 131–140. doi:10.1109/VR.2016.7504696. 2

[PZD*18] PRICE T. W., ZHI R., DONG Y., LYTLE N., BARNES T.: The impact of data quantity and source on the quality of data-driven hints for programming. In *Artificial Intelligence in Education* (2018), Penstein Rosé C., Martínez-Maldonado R., Hoppe H. U., Luckin R., Mavrikis M., Porayska-Pomsta K., McLaren B., du Boulay B., (Eds.), Springer International Publishing, pp. 476–490. 7

[RBB*00] RIZZO A., BUCKWALTER J., BOWERLY T., VAN DER ZAAG C., HUMPHREY L., NEUMANN U., CHUA C., KYRIAKAKIS C., ROOYEN A., SISEMORE D.: The virtual classroom: A virtual reality environment for the assessment and rehabilitation of attention deficits. *CyberPsychology & Behavior 3* (06 2000). doi:10.1089/10949310050078940. 2

[RGB13] RONFARD R., GANDHI V., BOIRON L.: The Prose Storyboard Language: A Tool for Annotating and Directing Movies. In *2nd Workshop on Intelligent Cinematography and Editing part of Foundations of Digital Games - FDG 2013* (Chania, Crete, Greece, May 2013), Society for the Advancement of the Science of Digital Games. URL: https://hal.inria.fr/hal-00814216. 2

[SB02] SAYE J. W., BRUSH T.: Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Educational Technology Research and Development 50*, 3 (2002), 77–96. 2

[VGD16] VAUGHAN N., GABRYS B., DUBEY V.: An overview of self-adaptive technologies within virtual reality training. *Computer Science Review 22* (09 2016). doi:10.1016/j.cosrev.2016.09.001. 2

[VL14] VAN LIER L.: *Interaction in the language curriculum: Awareness, autonomy and authenticity*. Routledge, 2014. 2

[WPRC18] WU H.-Y., PALÙ F., RANON R., CHRISTIE M.: Thinking Like a Director: Film Editing Patterns for Virtual Cinematographic Storytelling. *ACM Transactions on Multimedia Computing, Communications and Applications 14*, 4 (2018), 1–23. URL: https://hal.inria.fr/hal-01950718, doi:10.1145/3241057. 3

[You98] YOUNGBLUT C.: *Educational uses of virtual reality technology.* Tech. rep., Institute for Defense Analysis Alexandria VA, 1998. 2