How to Design Virtual Video Production for Augmented Student Presentations

Jonas Collin and Thomas Westin
Stockholm University, Department of Computer and Systems Sciences, Kista, Sweden
jcoll@dsv.su.se
thomasw@dsv.su.se

Abstract: E-learning environments have been developed and used by teachers and learners for decades. However, it is well known that sending, recording or meeting online can have a lack of presence and immersion. Furthermore, the configuration of a studio environment typically depends on physical props and technologies, which can be time consuming and hard to use for teaching purposes where each session may need a different configuration. Virtual Video Production (VVP) is a relatively new technology that builds on advances in extended reality (XR), supported by game engines and computer-controlled camera equipment. Camera data (pan, tilt, zoom, position) can be sent to a virtual camera in the game engine. The scene can be rendered via a green screen or with large LED displays. This provides an immersive presence with virtual 3D objects positioned in the room. Light settings can be mixed into the scene with remote control of LED lights to be in sync with virtual lights. Thus, VVP opens many opportunities for more immersive e-learning experiences. The challenge is how to apply these opportunities that involve syncing of several technical components and layers, designed to be useful within limited course resources. The question in this paper is how VVP can be designed and set up in an easy way for teachers and students to use it for presentations in courses? This paper presents results from a course in Immersive Environments where students first developed mobile Augmented Reality app prototypes and then used VVP to present their final work together with the authors. The authors documented preparation, wrote instructions for students and observed with notes taken during recording sessions. The results show how VVP can be designed and set up for course presentations that goes beyond a plain video recording in a lecture room or at home, but also beyond what was previously possible in a video studio at the university. This includes e.g., technical setup, direction of students, synopsis, concepts and virtual 3D props. Finally, the authors draw conclusions of what challenges remain for future research and suggestions of how to overcome them.

Keywords: technology-enhanced learning, extended reality, method, immersion, video production

1. Introduction

Virtual video production (VVP) has its roots in machinima (machine-cinema) where computer games and game technologies are used to produce videos, but also traces back to the so called 'demoscene' in the eighties to show off what is feasible to do with graphics and music on computers. (Harwood & Grussi, 2021). This paper first gives a background to early and current research efforts in this field, before presenting our own work with findings based on implementing virtual video production techniques in a video studio on campus. Finally, these findings are compared to the previous research, draws conclusions and presents suggestions for the future in this field.

1.1 Background

Early research papers in this field include the work of Lindley & Vercoustre (1998) and Lindley (1999) who used Standard Generalised Markup Language, SGML (a markup language similar to HTML) to synthesise videos based on an online video database, and a tree structured search (Lindley & Vercoustre, 1999). Lindley (2001) also developed an association engine for creating interactive videos.

Further early work was described in a technical report by Guil & Zapata (1998) who outlined issues of real time VVP such as scene rendering speed, syncing with the real camera parameters (lens settings, pan, tilt) and movement (3D-space coordinates). Guil et al (ibid) aimed to "develop a fast image processing algorithm to detect the 3-D camera position and the tilt and pan angles" based on a template placed on the camera, and then use other cameras to get tilt and pan from the angular distortion of the template. They mention mechanical decoders and vision techniques as alternatives. (Ibid.). However, electromagnetic, high-precision motion trackers that could be attached to various devices were also available at that time e.g., Polhemus Fastrak (Pimentel & Teixeira, 1993).

Spielmann et al (2018) presents an interesting collaborative production pipeline tool, built with Unity and Katana, which gets camera tracking data, light capture and estimation, and controls light settings in the room via a lighting server and DMX controllers. (A DMX or Digital Multiplex is the standard to remotely control lighting fixtures). The tablets provide an AR viewport into the virtual scene and can provide data for syncing the virtual and real world with anchor points. As an example of the technical fidelity needed to create the illusion of presence, James et al (2021) presents an approach for colour-management of LED walls used as background in virtual video production, enabling control of exposure and white-balance across the wall.

While the above technical opportunities and challenges seem fairly well researched during an extensive time period, research about the use of virtual production for teaching seems scarcer. However, Bennett & Kruse (2015) present pedagogical challenges and opportunities to teach visual storytelling via motion capture and visual effects in immersive environments. Recently, research has also been made by Nebeling et al (2021) to ease the use of virtual video production in education with compositing based on pre-sets, in their XRStudio where the production was made in virtual reality to make it more economically feasible without access to a studio environment with LED displays. In contrast, the study in this paper focuses on a real (actual) video studio and how to make it easier to use for education with virtual video production, also with a more affordable approach than LED walls. While adding LED walls provides better immersion for the presenters and easier lighting, many of the benefits of VVP can still be achieved without such expensive LED displays.

1.2 Challenge and research question

The challenge is how to apply the opportunities of VVP that involve syncing of several technical components and layers, designed to be useful within limited course resources. The question is how VVP can be designed and set up in an easy way for teachers and students to use it for presentations in courses?

2. Methods

This study takes a design science research approach (Johannesson & Perjons, 2021) by designing a virtual setup in a studio, with focus on explicating problems (challenges and opportunities). The methods were based on auto-ethnographies of a media producer (the main author) and a teacher in immersive environments (second author), reflecting on our own experiences from concept to production before, during, and after a course about immersive environments running for ten weeks in 2022. The experiences were based on both designing the virtual set, performing technical tests of equipment (cameras, lights, game engine, formats), as well as 3D-space positioning of cameras, and the possible area of recording with both virtual objects rendered in the studio, a background rendered on a large green screen as well as augmented reality (AR) apps developed by students. Instructions for a basic screenplay script with regards to virtual video production were also designed. Virtual production issues were observed while preparing for and then recording three student presentations during three three-hour sessions. The students have signed a consent form to participate with their faces blurred.

3. Findings

3.1 Physical studio considerations for VVP

This study was conducted in a video studio on campus enabling settings for immersion within limitations of what was available and affordable. This section provides details about how the studio was configured and designed, which can be helpful for others who want to create a similar setup and also understand how this study was done.

The floor layout, studio room (39 m2), control room and technical room, was designed by the main author in cooperation with interior architects and a professional audio engineer. In VVP, green screens are today typically replaced by large LED displays that enable actors to be partially surrounded by and see the virtual set directly and also lighting to be applied by the displays themselves. This is a limitation in this study where we did not have access to such displays, but this is also a more affordable variant of VVP for educational purposes. As an example, the main author built a green screen cyclorama (a large concave backdrop positioned at the back of the scene) with a ~€600 budget in 2016. It covers two walls and the floor. It was made of a green needle felt mat on a wooden framework. The needle felt mat was on sale on 3x30-meter rolls, which works both on the walls as well as on the floor where some other type of fabric could make you slip. See Fig. 1.

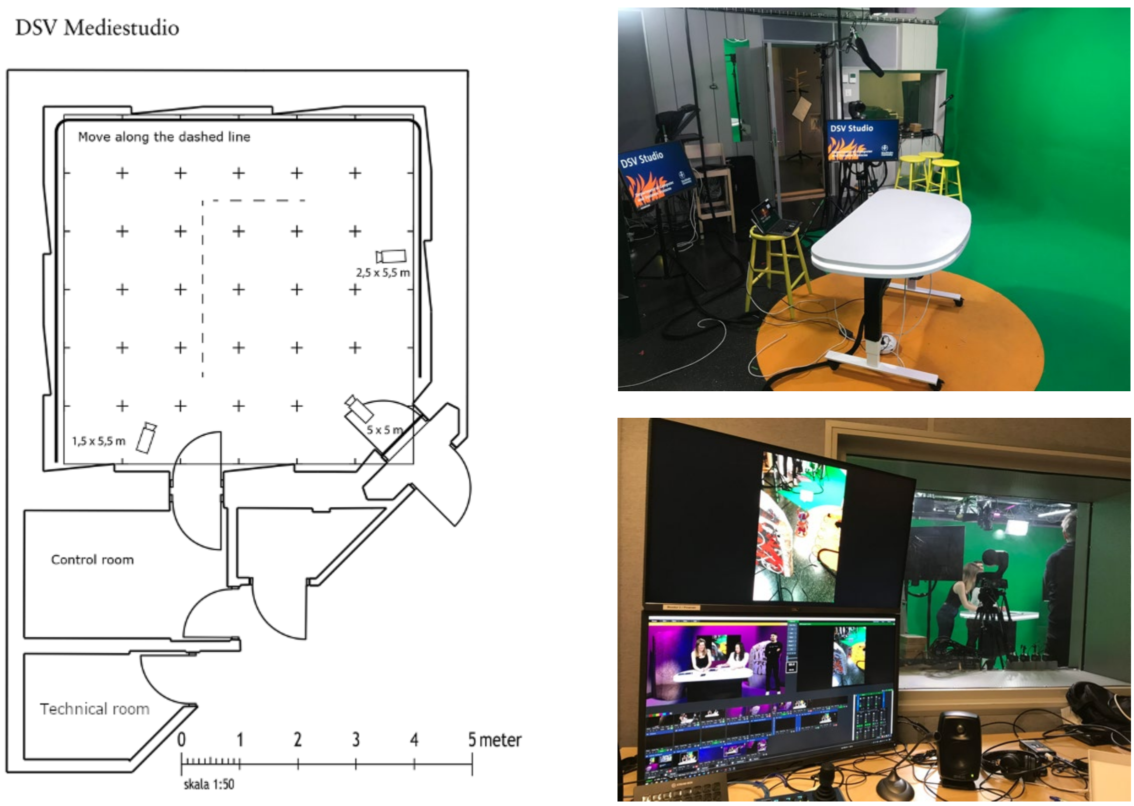

Figure 1: A map of the studio (left), the studio seen towards the control room with reference displays (top right), and the studio seen from the control room (bottom right)

The lighting was standard three-point lighting designed for three-camera shooting with LED-spotlights and fluorescent tubes for lighting the green screen. Recently, 12 RGB-tubes (red, green, blue) for ambient, coloured/white light set up in a grid in the ceiling were added. The lighting can be controlled by a USB-to-DMX-interface hooked up to a laptop with lighting control software. DMX can be controlled by the Unreal Engine game development software, which makes it possible to map the real studio lighting to the virtual world. We have not implemented DMX in the Unreal engine yet, but we did use the RGB-tubes to make the illusions that the virtual walls reflected their light on the students (Fig.2 and Fig.3).

Figure 2: A virtual video production view with background, synced lighting, a virtual presentation screen (here showing live content of an AR app from a tablet) and virtual 3D-objects (in this case, rune stones)

The table had a half-moon shape with an RGB LED strip on the edge (Fig.2). The shape makes the actors open up their bodies facing the camera, thus prohibiting them to be in profile to be more inviting to the audience. The LED strip gives a TV production style. We believe that a studio table with a non-standard appearance fits better in a (virtual) production. Photoshop and free image archives were used to experiment with green screen backdrops in 3-camera setups, to check eye levels of participants and focal length on the camera as it works good enough for fast studio productions. Two camera angles are illustrated in Fig.2 and Fig.3.

We used shotgun microphones (for directed audio recordings with good pickup of voices and little of noise in the room) on boom stands and wireless lavalier (wearable) microphones. We used two Panasonic UE150 PTZ cameras (professional cameras with pan, tilt, zoom that can be remotely controlled with a joystick) and a Panasonic EVA1. We used an HP Z8 workstation with an NVIDIA Quadro RTX 5000 graphics card and BMD Decklink 8k capture card. The cameras were connected to the capture card with SDI-cables (for full dynamic reproduction). We used vMix as video mixer software as it was already configured in the studio but Open Broadcaster Software (OBS) could also have been used. The idea behind using a Windows machine with software instead of using a studio-in-a-box solution like Tricaster is that we wanted to be able to experiment with all sorts of software for production in different settings.

In the studio there were three big monitors on stands showing the recording video (two of the visible in Fig.2). These could be placed so that the actors could see themselves when they interacted with virtual objects on the green screen, for example pointing at a virtual PowerPoint presentation. The monitors could also be used for showing the desktop of the studio workstation to mirror the controls of the control room, enabling an actor or teacher to control everything from within the studio room. Furthermore, the camera joystick and the video mixer control panel could be used in the studio room as well.

Figure 3: A close-up of the scene from a different camera, more clearly showing highlights in the hair and clothes that is synced with the virtual light on the runestone, to create an integrated and more immersive experience

3.2 Implementing Unreal virtual production in the studio

Our Panasonic UE150 cameras come with built in camera tracking. They can broadcast the pan, zoom, tilt and lens data to third party software using the Free-D protocol but cannot track their position in the room. Thus, we started by measuring the studio floor placing out the cameras at exact coordinates in the x, y and z axes. We installed Unreal at the studio workstation. In Unreal, we created a simple virtual studio in the same scale as the studio and placed virtual cameras at the same positions as in the real-world studio.

We used the Disrupt AR plugin in Unreal to read the Free-D tracking data from the cameras and the Live Link plugin to map the tracking data to the virtual cameras. This allowed real-world cameras to direct the pan, tilt, zoom and focus of the virtual cameras. The third camera (the Panasonic EVA1) did not have a tracker and was treated as a static long shot in both worlds.

The images from the virtual cameras were sent to the video mixer (vMix) on the same computer to use as green screen backgrounds. This created the illusion of mixed reality. To feed the video from Unreal to vMix we used the Network Device Interface (NDI) plugin in Unreal. We also used NDI to show the PowerPoint within the Unreal virtual studio. All these settings could be found in several tutorials online.

In vMix you can delay video and audio streams and thus sync all media exactly. The problem with NDI was that there was unpredictable lag. We guess it has to do with its extensible use of the GPU when compressing data via NDI (packing/unpacking) and processing the network data to display data. A generator lock (genlock) should have been used to really sync all audio and video sources.

3.3 Preproduction instructions to students

The students first made a 20 minutes visit in the studio per group to get an idea of possibilities and limitations of the studio. They also got written instructions of what to think about, in terms of prerequisites and a synopsis.

The studio prerequisites which the students were informed about were: *1) A large green screen cover the top and right walls in Figure 1; 2) The three studio cameras cannot be moved while recording as their position in the room cannot be tracked and have to be synced manually with the virtual studio in Unreal; two cameras have remote control of pan, tilt, zoom; 3) Shoot towards the green screen wall closest to the control room* (to ease the setup between groups); *4) 3D-objects in Unreal will not be rendered correctly in depth, which means you cannot go behind as this will still look as you are in front of the 3D-objects; 5) Lighting is possible with colour LED-panels and tubes* (which was also prepared beforehand)*; 6) Three reference displays are available that can be repositioned so you see what you are doing in relation to the green screen; 7) A basic predefined virtual studio environment is used, but your own 3D-objects made with photogrammetry will be added a day before the recording day; 8) AR tablets are connected via a USB-C cable to HDMI; and 9)) The light setup and camera focus are fixed which means you must only move along the dashed line in Figure 1.*

The students also got instructions of writing a basic synopsis to think about: *1) Do you want to have a table or not in the scene, and approximately where in the room?; 2) where do you want to position your AR objects with the tablet?; 3) How do you want to cut between camera sources, and approximately when while recording? Or do you want to edit and cut in post-production? Sources can be three studio cameras, the AR-tablet app and another NDI source (e.g., a Powerpoint).* The synopsis details about tables and positions were helpful to plan the recording session and know where to position cameras. The students also got a link about virtual production in Unreal (https://www.unrealengine.com/en-US/virtual-production) if they wanted to learn more. The students' 3D-objects were based on their own photogrammetry (3D-scanning) of real rune stones from the Swedish History Museum. For some unclear reason, importing high resolution versions of the objects into the scene resulted in distorted models, whereas a low polygon version worked fine (as seen in Fig.2 and Fig.3). Due to limited time this issue was not resolved before the recording session.

Both the studio visits and the instructions were crucial as the students had not worked with virtual video production before. As can be seen from the above detailed description, the studio is a complicated technical environment allowing for experiments but also failures if you do not reflect on how to use it. They got two hours for recording sessions in the studio, one per group of approximately five students. This limited time also called for preparations so videos could be recorded for all groups of students, by the course manager of what content could be included, but also for the media producer to make pre-sets and syncing ready before the recording sessions.

3.4 Production direction of students

When the students arrived on the recording day, they were asked to explain their desired outcomes and their thoughts on how this could be achieved. We helped them to adapt their ideas to the studio setting and gave them directions on how to act into the camera. We also discussed the basic dramatic structure of their presentation. We made sure that they had a beginning, middle and an end. They got some directions on which camera to look into and that they had to use more energy in their acting. These were media students so they already had some basic training on how to act in front of a camera.

We rehearsed a couple of times. Then we shot the presentation in one take. We recorded all cameras and the live cut version in separate video clips. We transferred all media to the students' own media space. The students were free to use the live cut or to edit their own version from all the media clips. We noticed a need to move the cameras while recording but as everything needed to be in sync, we did not do that.

Some special considerations in the studio regarding presenting augmented reality (AR) apps that the students had made in the course they attended was that: 1) Placing AR objects on walls with green screens required us to cover part of the right wall with a curtain, since AR requires a varied surface for tracking to work properly; 2) Scene tracking is a 3D-scanning technique for positioning objects in the room relative to an anchor and also enables multi-user AR by sharing anchors via the cloud. However, scene tracking requires pre-scanning of the room with the camera where the room setup cannot change too much in-between, and also requires a fairly large open space to make it work.

4. Discussion

To answer the research question, it is useful to make comparisons of our study with the related research. The early work by Lindley & Vercoustre (1998, 1999) resembles some of the early unpublished work made by the main author of this paper, unknowingly of each other at the time. Challenges found in related research were about rendering speed, synchronisation of virtual and real cameras, lighting, and colour management as well as course specific challenges. Some opportunities were co-location/remote teaching, collaborative production, compositing based on presets and using AR to enable sensor data for scene management.

While many of these challenges and opportunities found in the earlier studies remain today, the latest technologies enable new possibilities for teachers and students for more immersive learning experiences. This is further enhanced by free software (e.g. Unreal and OBS) that works with both consumer level hardware as well as in professional studios. However, how to involve teachers and students in an efficient way is not trivial, especially if they have little time, experience of or interest in VVP in itself. In this study, both the involved students and teachers had a focus on immersive environments, so the need for an easy production pipeline can be assumed to be even greater in other types of courses.

4.1 Challenges and opportunities with VVP in education

As found in the related research, VVP has been around as a concept at least since 1998, long before large LED displays existed. While large LED displays are the standard in virtual movie productions today, green screens and reference displays can be used as an affordable option for VVP in education. Large green screens can be made with a low budget. Unreal with plugins is a feasible way to create educational VVP content, also with a low budget except for what is available in a typical video studio. Both approaches allow that the studio model is transferred to different 3D worlds to change location without the need to change the rig of the studio in between educational activities and courses.

A central part of the technical setup is to build a virtual studio environment in Unreal with the same scale as the actual (physical) video studio. One challenge with this was that the virtual and real cameras had different scaling and it was difficult to know what settings to change. While the technical considerations of the studio environment and implementation in Unreal is complex and requires technical expertise to set up, the VVP recording sessions went smoothly and without any severe issues that could not be resolved easily within two hours for each group of students. Outlining prerequisites and providing a visit for students before they planned their presentations with a synopsis helped all stakeholders involved (media producer, teacher, students). This shows the potential for this approach, but the limitation is today the need for involving a media producer that can handle the technical setup (lighting, positioning, audio and more).

4.2 Comparing with VVP research for teaching and learning

Going forward we can learn from the requirements that Nebeling et al (2021) propose for production: 1) “High-quality video without a production team.”; 2) “Automating the production process”; and 3) ““No-technology” solution for instructors". In this study the first requirement was met with video quality but required an experienced media producer to be present. The second requirement was not fulfilled, as the media producer had to manually set up the technical environment. The third requirement was fulfilled. Nibeling et al (ibid.) also propose instructional requirements, but those focus on VR and not studio production as we do in this study.

Nebeling et al (2021) present a walkthrough of XRStudio with similarities and differences to our approach. The main difference is that they did all production in the virtual XRStudio while we used an actual video studio, but they also provided tools for 2D and 3D sketching and manipulation. There were also several similarities: 1) presenting a default scene with pre-calibrated cameras, which the media producer had prepared in the studio; 2) they had pre-sets for live production with different sources (cameras, slides and more), which we had in the control room; 3) they had mixed reality capture, that we achieved with both studio cameras and the green screen with 3D objects, as well as with the AR apps developed by the students; 4) record and replay, which is naturally possible in a video studio; and 5) they provided multiple output modalities for live or recorded streams, that we also have in the studio, but we only used it for recordings in this study. Also, as Nebeling et al (2021) points out, there are challenges in relying on instructors only to produce videos; e.g. using transitions without thinking about the audience of students is a good example. In this regard, education of using either a VVP studio or a virtual XRStudio is required or involving a professional media producer to some extent.

4.3 A proposed setup for teachers and students

To be useful for students and teachers the studio environment needs to be preconfigured with pre-sets, selected via an easy-to-use interface. This can be achieved with physical, programmable buttons or with similar digital buttons in an app. The pre-sets include light settings, green screen background and 3D-objects and virtual cameras synced with coordinates in the physical studio, and virtual screen to display presentation content. With these pre-sets, students and teachers could simply use the remote-control joystick for camera pan, tilt, zoom need with some instruction. OBS studio should be used instead of vMix so teachers and students can practice with the applications also when not having access to the studio, along with Unreal as both are free software.

By using pre-sets for each course and reducing the complexity to one joystick control and not trying to cut between cameras live, it should be possible to let students and teachers to use virtual video production without much intervention by a media producer in recording sessions. If the joystick is positioned in the studio instead of the control room, it can also simplify the situation further. It may be possible to avoid using the control room altogether, in line with the "no technology" approach. However, not cutting live means that there may be need to make some editing after the recording session, which requires some extra time and basic skills in video editing software. Thus, it may be a better approach to have one person to cut between cameras live, and one person that handles the joystick to control pan, tilt and zoom.

5. Conclusion and future work

The research question was: How can VVP be designed and set up in an easy way for teachers and students to use it for presentations in courses? To answer the research question, we have described our own experiences from both designing and preparing a virtual video production set, and observations while students produced short presentations in a video studio with virtual video production. Based on this we have discussed how VVP can be designed and set up for teachers and students with suggested instructions of prerequisites, questions for writing a synopsis, and what direction was needed during the recording session.

5.1 Future work

In our future work we aim to conduct a study where other teachers and students are involved in the entire design process to learn more about how different courses can utilise the potential of virtual video production for teaching and learning. When implementing a new technology an aim is to not disturb the existing workflow. We would like to create some kind of template or application that can compose the video streams and feed them into our video mixer. When the app starts it should just work. If you want to change the environment or other features there should be a graphical interface for this. We would like the cameras to report their position in the room so that we do not have to measure manually. This could be fixed by implementing a tracking system like HTC Vive or using their new HTC Vive Mars System. We must also deal with the lag and sync problems. This could be fixed by running the Unreal engine on a separate computer and also sync the cameras with genlock.

We have not explored the use of this technology fully in this first explorative attempt. It could be great in subject areas like history to let the teacher be a guide on an ancient site, or in case-based learning where the teachers can act in different environments creating more immersive experiences or in complex technical environments where you can show how it works. Further, it could be used just to create instructional videos that look nice.

With the three cameras positioned in the studio and within the virtual world it is easy to move the whole studio around to different locations without having to do any rigging in the studio. Thus, it could be really time efficient. One possibility would be to try and automate some of the setup and provide AR-based support for what to do (and how) in the studio. The media producer may then be less of a technical support person, and can focus more on direction of teachers and students in acting in front of a camera.

References

Bennett, G. & Kruse, J. 2015. Teaching visual storytelling for virtual production pipelines incorporating motion capture and visual effects. SIGGRAPH Asia 2015 Symposium on Education. Kobe, Japan: Association for Computing Machinery.

Guil, N. & Zapata, E. L. 1998. Real-Time Virtual Video Production. Malaga, Spain: University of Malaga.

Harwood, T. G. & Grussi, B. 2021. Pioneers in Machinima: The Grassroots of Virtual Production, Vernon Press.

James, O., Achard, R., Bird, J. & Cooper, S. 2021. Colour-Managed LED Walls for Virtual Production. ACM SIGGRAPH 2021 Talks. Virtual Event, USA: Association for Computing Machinery.

Johannesson, P. & Perjons, E. 2021. An Introduction to Design Science, Springer, Cham.

Lindley, C. A. & Vercoustre, A. M. 1998. Intelligent Video Synthesis Using Virtual Video Prescriptions. International Conference on Computational Intelligence and Multimedia Applications. Gippsland, Australia: River Edge.

Lindley, C. A. Generic film forms for dynamic virtual video synthesis. Proceedings IEEE International Conference on Multimedia Computing and Systems, 7-11 June 1999 1999. 97-101 vol.2.

Lindley, C. A. & Vercoustre, A.-M. Generic Viewer Interaction Semantics for Dynamic Virtual Video Synthesis. In: Huijsmans, D. P. & Smeulders, A. W. M., eds. Visual Information and Information Systems, 1999// 1999 Berlin, Heidelberg. Springer Berlin Heidelberg, 83-90.

Lindley, C. A. 2001. A Video Annotation Methodology for Interactive Video Sequence Generation. In: Earnshaw, R. & Vince, J. (eds.) Digital Content Creation. London: Springer London.

Nebeling, M., Rajaram, S., Wu, L., Cheng, Y. & Herskovitz, J. 2021. XRStudio: A Virtual Production and Live Streaming System for Immersive Instructional Experiences. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Yokohama, Japan: Association for Computing Machinery.

Pimentel, K. & Teixeira, K. 1993. Virtual Reality: through the new looking glass, USA, Windcrest Books.

Spielmann, S., Helzle, V., Schuster, A., Trottnow, J., Götz, K. & Rohr, P. 2018. VPET: virtual production editing tools. ACM SIGGRAPH 2018 Emerging Technologies. Vancouver, British Columbia, Canada: Association for Computing Machinery.