Design and Prototype Implementation of MR Pre-Visualization Workflow

Ryuhei Tenmoku* Ryosuke Ichikari† Fumihisa Shibata† Asako Kimura† Hideyuki Tamura‡

*Research Organization of Science and Engineering, Ritsumeikan University
† ‡Graduate School of Science and Engineering, Ritsumeikan University

Abstract

This paper describes a pre-visualization system using mixed reality (MR) techniques for filmmaking. This system is a prototyping production of the "MR-PreViz Project," was conducted performed in Japan. In the pre-production process of filmmaking, PreViz, pre-visualizing the desired scene by CGIs, is used as a new technique. As it is an advanced approach, we propose MR-PreViz that utilized mixed reality technology in current PreViz. MR-PreViz makes it possible to merge the real background and the computer-generated humans and creatures in an open set or at an outdoor location. The user can consider camera-work and camera blocking efficiently using MR-PreViz.

Keywords: Mixed reality, filmmaking, pre-visualization.

Index Terms: H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities—

1 Introduction

The Key Technology Research Project on Mixed Reality Systems (MR-Project) carried out in Japan from 1997 to 2001 gave significant impact on works of augmented reality (AR) and mixed reality (MR) [1, 2]. In general, people ask us "what are you going to do next in Japan" at annual ISMAR symposium. People expect us to open a new and unique frontier as the ARVIKA Project [3], which specializes its work on industrial applications. The MR-PreViz Project described in this paper is one of our answers to such expectation. In this project, we have tried to change the style of filmmaking using MR technologies.

Several works already exist on the application of MR technology to filmmaking. One such application is "2001: An MR-Space Odyssey" [4] developed by the MR-Project. This system enables for actors to confirm composition of live action and computer generated images (CGIs), which is done at the post-production stage in ordinary filmmaking process, in movie studio by wearing HMDs. One of the recent approaches is to make an interactive storyboard at the preproduction stage [5].

We promoted the five-years-project, "MR-PreViz Project" since Oct. 2005. This project is forwarded by Ritsumeikan University, Nara Institute Science and Technology, and Kyoto University. The central location of these institutes is Kyoto, which is the heart of Japanese filmmaking, especially, samurai movies. The purpose of this project is to develop authoring tools applicable to the entire filmmaking industry using MR techniques, which can merge real and virtual worlds [6, 7]. Thus, realizing a system that helps effective shooting in the production environment by providing pre-visualized images in which predefined CGIs are superimposed onto the actual scene even in an open set or location site.

*e-mail: tenmoku@rm.is.ritsumei.ac.jp
†e-mail: {ichikari, shibata, asa}@rm.is.ritsumei.ac.jp
‡e-mail: HideyTamura@acm.org

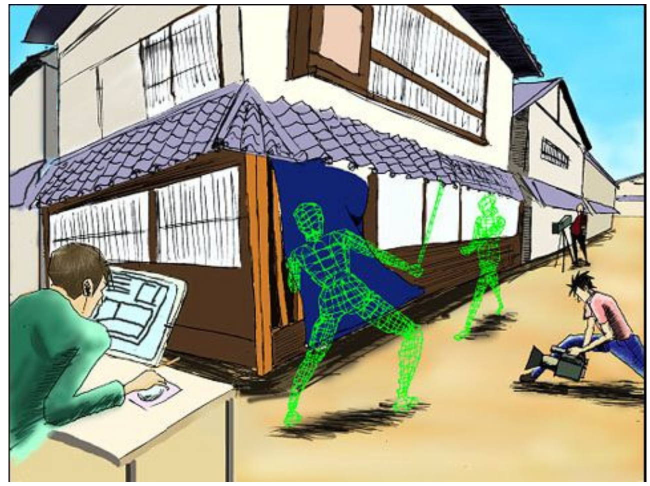


Figure 1: Conceptual illustration of MR-PreViz

This paper is structured as follows. Section 2 describes the basic concept of the MR-PreViz system. In Section 3, details of the MR-PreViz system are described. Sections 4 and 5 describe the workflow of the MR-PreViz and the state of the actual shooting of MR-PreViz movies, respectively. Finally, Section 6 summarizes the present work and also gives the future work.

2 Basic Concept

Recent filmmaking process has the following three stages:

Preproduction stage:
: Planning, scripting, casting, location hunting, writing storyboards, and constructing props

Production stage:
: Rehearsals, actual shooting, and special effects

Postproduction stage:
: Film editing, sound mixing, and visual effects (VFX)

In the preproduction stage, storyboards have been traditionally used for the interpretation of the director's intention by the staff. However, it is not difficult to imagine the limitation of storyboards for smooth image interpretation. Recently, pre-visualization (PreViz) [8], which is sometimes called Animatics, has been used to get to the further stage of the storyboards. PreViz is the visualizing technique based on computer generated images for discussing action scenes, camera angles, camera blockings, lighting conditions, and other situations and conditions before the actual shoot. Recent epic films have often used PreViz technology, and some of these films visualize whole scenes in advance using computer graphics. PreViz enables the director to effectively share thoughts with actors and staff. Current PreViz movies are depicted by relatively simple CGIs, since their purpose is to show positions and motions of

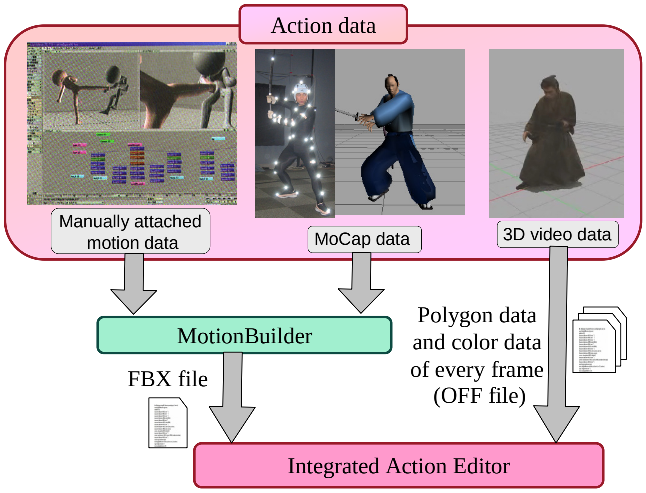


Figure 2: Input data of Integrated action editor

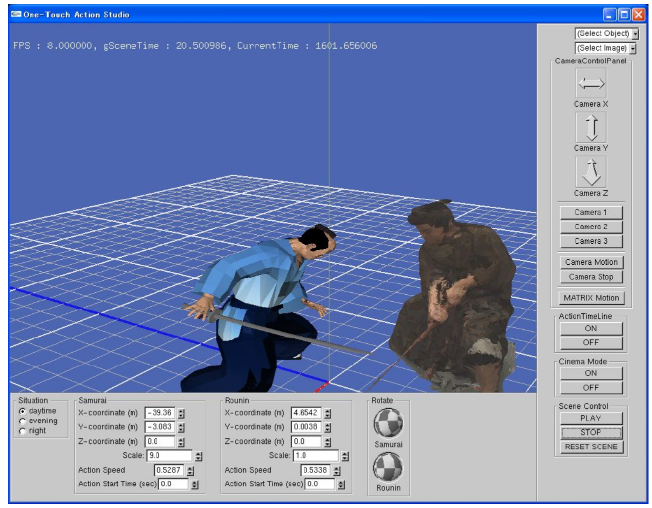


Figure 3: Example image of Integrated action editor

humans and creatures in a particular scene. However, fully computer generated images have innate limitation in visual reality. A perfectionist director may expect higher level of PreViz images. It is quite natural that she/he would want to confirm if the imagined picture compositions are suitable for the actual scene, especially, when a large open set or overseas location is selected.

In the MR-PreViz system, pre-visualizing movies are generated for sharing images of the final version and considering camera works and camera blocking. The target of our MR-PreViz system is not only scenes where computer generated creatures or vehicles play some role but also live action scenes where actors play her/his role. One can capture motions of a stunt man who performs blistering actions instead of the actors. This helps in effective preparation for the actual shooting. Figure 1 shows the example image of shooting the MR-PreViz movie.

3 Architecture

The MR-PreViz system consists of the following three software tools and a hardware system.

(a) Integrated action editor
We can edit actions of CG characters and creatures of various data types.

(b) 3D-space layout tool
Virtual and real objects are arranged using this tool to construct the MR-PreViz scenes.

(c) Camera-work authoring tools
They are tools for examining camera-work and camera blocking.

Detailed explanations for software tools and the hardware system are described below.

3.1 Integrated action editor

The integrated action editor is a tool used for preparing series of action sequences of CG characters and creatures to merge them with the real background. The users can easily construct blistering action sequences involving multiple actors and/or creatures. Our integrated action editor is designed to handle multiple motion data functions for editing motion data.

3.1.1 Handling multiple motion data

The dataflow of the integrated action editor is shown in Figure 2. As described in this figure, this tool can handle three types of motion data: manually inputted CG animation data, motion captured (MoCap) data, and 3D Video data. MoCap data is attached to the CG character using MotionBuilder (Autodesk Corp.). Since the output data type of MotionBuilder is FBX data, manually attached motion data is also inputted into this tool as FBX data. “3D Video” is a technology that allows one to reconstruct an image seen from any viewpoint in real time from video images taken by multiple cameras [6]. It is a kind of video-based rendering and is sometimes called “Virtualized Reality” [7]. We adopted the method developed by the Kyoto University [8]. 3D video data is inputted as polygon data and color information of every point as the OFF data type.

Three kinds of action data can be drawn using OpenGL or DirectX libraries because all are polygon-based data type. However, to handle different file types of action data, we face another problem of the difference in the sampling rates of these data. The sampling rate of current 3D video data is lower than that of the other types of motion data. We solved this problem by the nearest-neighbor approximation on the time axis.

3.1.2 Functions for editing motion data

Figure 3 shows an example screenshot of the integrated action editor using graphical user interface (GUI). To construct action sequences by editing three kinds of data, this tool has the following functions:

- Setting positioning and posture:
  Setting the initial position and posture of 3D model data freely.

- Scaling:
  Setting the scale of each action data.

- Setting playback speed and timing:
  Controlling playback speed and timing of each action data.

In addition, users can control the position and orientation of the viewpoint, pause and play the animation data, and display guiding grids or coordinate axes.

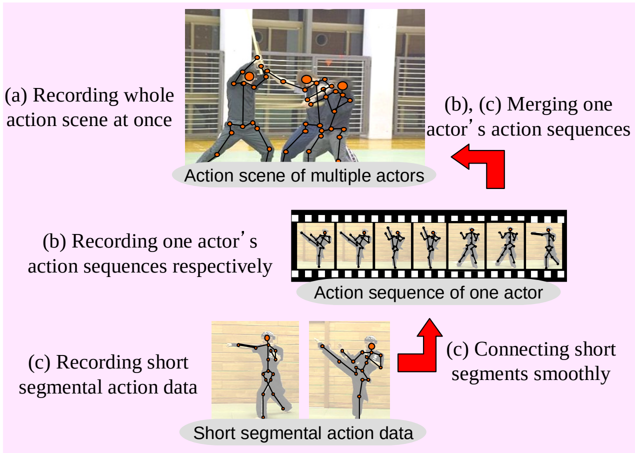


Figure 4: Conceptual image of archiving action data

3.1.3 Archiving action data

We have three concepts for constructing action scenes involving multiple actors and archiving action data.

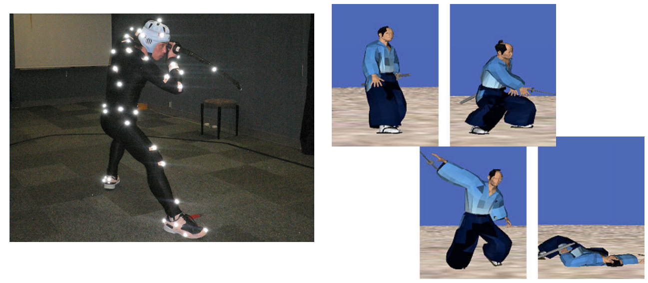
(a) The assumed whole action scene is recorded at once.

(b) The action scene is constructed by merging one actor's action sequences data recorded respectively.

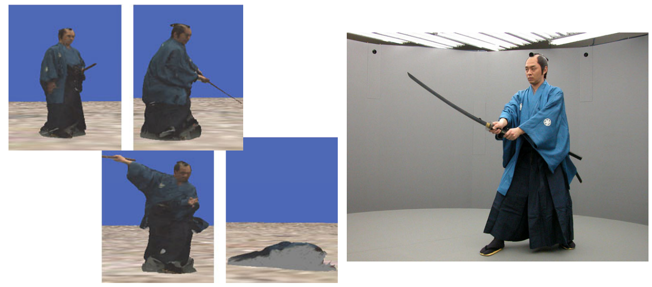
(c) First, various short segmental action data of one actor are recorded. Second, one actor's action sequence is constructed by connecting these short action segments smoothly. Finally, by merging one actor's action sequences, we obtain multiple actors' action scene.

Figure 4 shows a conceptual image of these methods. Method (a) does not have any paradoxes in timing of actions and actor's positions, because recorded actions sequences are used without any handling in this method. However, in this method, there are some problems. If the director wants to change the action scene slightly, the entire action scene has to be recorded again. Moreover, occlusions occur in recording action scenes. The method (b) does not have the occlusion problem. Additionally, we can easily change the positional relationship of actors in this method. However, we need to verify whether we can construct well-timed action scenes when combining the action sequences that have been recorded individually. According to us, the concept of method (c) realizes that infinite action scenes are built from finite archived action data. However, we need to discuss whether the action sequences can be segmented into short segments. In addition, this method of connecting action data has many problems.

Thus, we examined the availability of methods (b) and (c). In the first validation, we examined the validity of method (b). We constructed a Chanbara scene with two samurais in consultation with a professional swordfight arranger and recorded this scene as one actor's action sequences, which were played by two professional samurai actors. We acquired these sequences as MoCap data and 3D video data. Figures 5 shows data acquisition scenes and examples sequences of MoCap and 3D video data. We have reconstructed four combinations of two samurais' battle (MoCap vs. MoCap, MoCap vs. 3D video, 3D video vs. MoCap, and 3D video vs. 3D video) by merging two continuous action sequences using our integrated action editor. We constructed a consistent battle scene with good timing. This shows that performances of professional actors are precise enough to confirm the availability of archiving action data individually to be merged.



(a) MoCap data



(b) 3D video data

Figure 5: Data acquisition scenes and examples of action data

Next, we tried to examine the feasibility of method (c). The traditional Japanese style of using swords (Chanbara) and the Chinese kung-fu action have certain rules called "Kata". Accordingly, we archived action sequences that were divided into fundamental action data based on "Kata". Every fundamental action data starts and finishes with the same pose (basic fighting stance). In addition, we tried reconstructing one samurai's continuous action sequences by connecting recorded fundamental action data. As the first step of this method, we play the recorded fundamental action data continuously. However, in this simple method there are some displacements of CG actors in boundary and redundant parts between fundamental action data. In other words, this area has some research issues, and we try to resolve them for improving our MR-PreViz.

3.2 3D-space layout tool

3D-space layout tool realizes to design MR spaces by arranging CG objects, real background images, and CG action data, which is edited by the integrated action editor. This tool is used with the purpose of designing a virtual space simulating the actual shooting location. Although some 3D CG modeling OTS software are available for constructing CG-based 3D spaces, this is the first trial that constructed 3D spaces that could be applied MR.

The outline of the 3D-space layout tool including dataflow is shown in Figure 6. The users of this tool roughly design 3D spaces which simulate actual shooting site beforehand. When the MR-PreViz movie is shot, users need to align real and virtual world precisely. Our 3D-space layout tool is constructed by two sub tools: VR-space layout tool and MR-space layout tool. The former is used for the rough design of the shooting site in the virtual space and the later is used for the fine adjustment in the location site.

We designed the VR-space layout tool supporting the following layout functions:

- Arrangement, movement, rotation, and scaling of CG objects

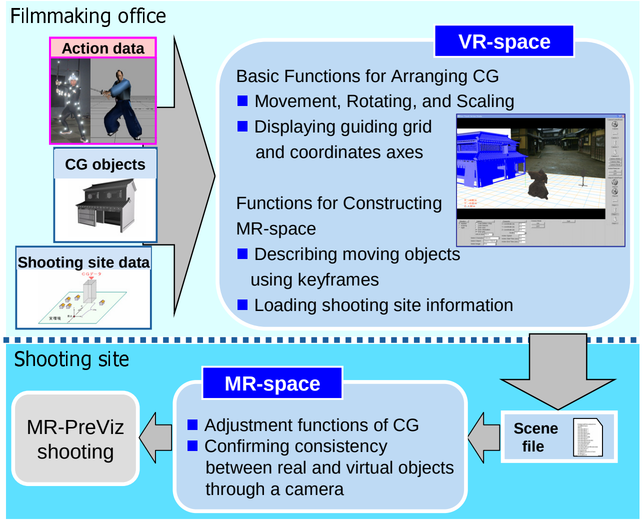


Figure 6: Outline of 3D-space layout tool

Basic data	Trim data	
Scene number	trim 1	File name of sequence
Location site		Camera-work
Equipment info. (camera, lens, lights)		Camera blocking
Aspect ratio	trim 2	File name of sequence
Script of the scene		Camera-work
Storyboard of the scene		Camera blocking
Output files of other tools	⋮	
	Order of connecting trims	

Table 1: Data structure of CWML

- Loading shooting site information to realize easy layout
- Outputting scene graphs of the designed space
- Specifying keyframes for describing movements of CG structures other than characters

The VR-space layout tool also has functions of displaying guiding grids or coordinate axes. Additionally, the MR-space layout tool has the adjustment function between real and virtual worlds. The user of MR-space layout tool can adjust real and virtual objects by viewing generated MR images.

3.3 Camera-work authoring tools

The "camera-work authoring tools" is the collective term of tools that support examining the camera-work and camera blocking [11]. Camera-work authoring tools consists of the camera-work recorder and the MRP browser. The camera-work recorder is loaded into the MR-PreViz image capturing and compositing system described in Section 3.4. The dataflow of camera-work authoring tools is shown in Figure 7. MR-PreViz movies are recorded with camera-work and camera blocking information. They are described in Camera-Work Markup Language (CWML). Recorded CWML data is utilized in actual shooting by interpreting with an MRP browser.

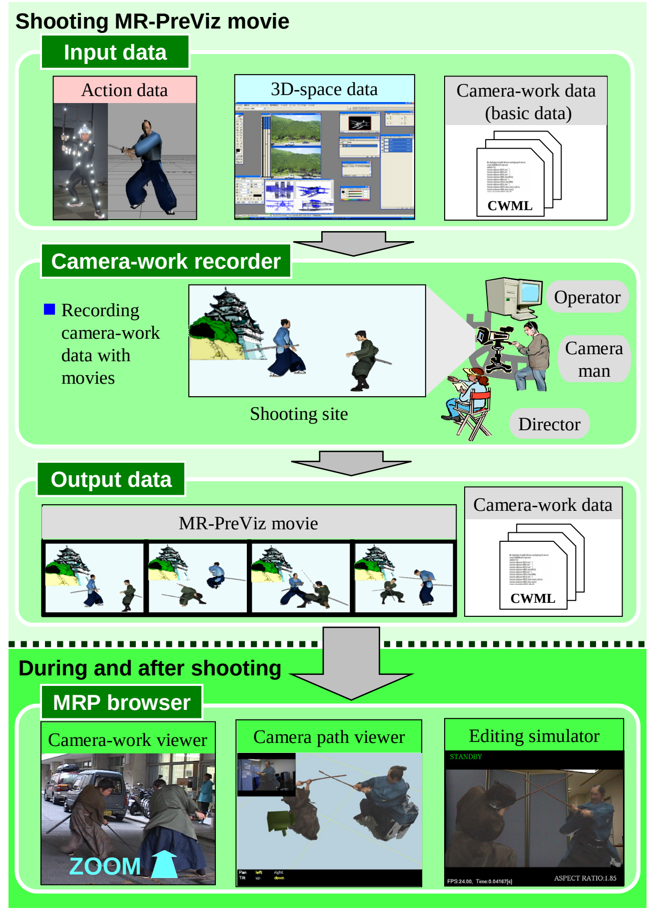


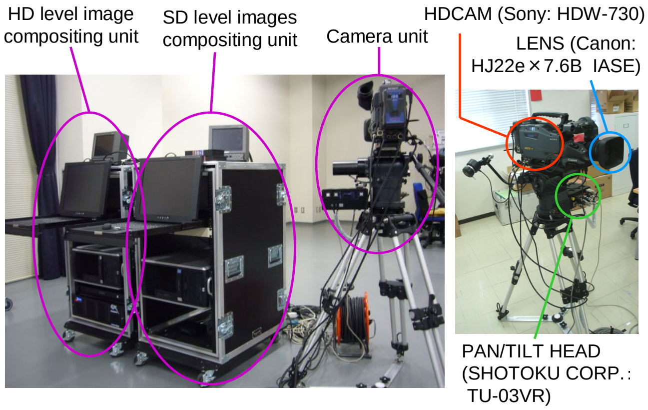
Figure 7: Dataflow of camera-work authoring tools

3.3.1 CWML

CWML is a new description language for describing camera-work as an instance of XML meta-language (XML document). Table 1 shows the data structure of CWML data. CWML data is composed of the basic data and the trim data. The basic data includes scene information, equipments information, and so on. All of the basic data have to be inputted manually before shooting MR-PreViz movies. On the other hand, the trim data includes camera-work data, camera blocking data, and file names of movies. In our MR-PreViz project, the word "camera-work" is defined as the camera position, the posture, and the zooming factor of every frame during the shooting. Camera-work data is recorded automatically during the shooting MR-PreViz movies.

3.3.2 MRP browser

The MRP browser is viewer software to visualize data described in CWML. The relation between CWML and the MRP browser is similar to that between HTML and Web browser. This browser is used after shooting MR-PreViz movies to check camera-work and camera blocking. We assume that the MRP browser is used for confirming the recorded camera-work in the location site of MR-PreViz movies, discussing the recorded camera-work in the office, and confirming the camera-work in the actual location site. In any cases, the movie crew can check and discuss the recorded camera-work and camera blocking as viewing visualized CWML data and captured MR images. The prototype of our MRP browser is designed to have the following viewer functions:



(a) MR-PreViz Image capturing and compositing system (b) Camera unit

Figure 8: Image capturing and compositing system

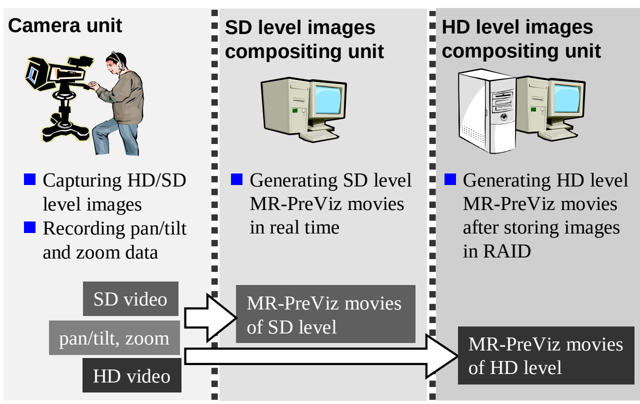


Figure 9: Workflow and dataflow of prototype hardware system

- Camera-work viewer :
  The camera-work viewer function superimposes the recorded camera-work information in easy-to-read format as shown in the example images of Figure 7. We assume that this function is mainly used for confirming the camera-work in the actual location site.
- Camera path viewer :
  The recorded camera path is shown in CG-based location site with camera parameter information.
- Editing simulator :
  The editing simulator function allows the user to simulate editing movies as viewing the recorded MR-PreViz movies. This function supports specifying editing points and the order of connecting trims. The MRP browser also can indicate editing points of recorded MR-PreViz movies and camera blocking information.

According to us, both camera-work and camera blocking, as well as editing points have to be shown with recorded MR-PreViz movies to the user. We further need to discuss the use of these functions to construct the interfaces of this browser.

3.4 MR-PreViz image capturing and compositing system

A high definition (HD) digital camera is used for the MR-PreViz system even if films are used for the actual shooting. However, the purpose of MR-PreViz is to assist decision of various factors in shooting such as camera-works and picture compositions, lenses, lights, and other necessary equipments, or at least close to the traditional ones that are normally used in filmmaking. The current core system capturing and compositing of MR-PreViz movies is shown in Figure 8. Our system consists of camera unit and SD and HD level images compositing systems.

The workflow and dataflow of present image capturing and compositing system is shown in Figure 9. The camera platform has two rotary encoders that detect pan and tilt angle of the camera and transfer them as digital signals to the computer. Since the zoom lens of the camera has a functionality to detect the optical zoom factor, CGI composition with the actual scene is performed in real time according to the motion and zooming of the camera. However, because of the competence of the capturing board on PC, our first prototype can only composite and display movies of standard definition (SD) level in real time. It takes a while until the movies of HD level are produced. The camera posture of this system has only two degrees of freedom: panning and tilting. However, more degrees are required if you want to shoot using a dolly, crane, and other equipments. We plan to realize such an advanced system in the next stage of our project.

4 Workflow

In this section, the workflow of the filmmaking using MR-PreViz is described in detail. Figure 10 shows a part of the workflow of filmmaking using MR-PreViz. We have classified these processes into the following four phases: pre-award meeting, data arrangement, shooting of MR-PreViz, and actual shooting. Processes before making storyboards/CG-based PreViz movies and after the actual shooting are not, in particular, different from those of traditional filmmaking. Detail processes of each phase are described below.

[Phase 1: Pre-meeting]
First, the director and the staff select scenes suitable for MR-PreViz, for example, scenes including blistering actions or elaborate camera-work as seeing the storyboard or CG-based PreViz movie.

[Phase 2: Data arrangement]
After obtaining short segments of action data, the operator of the integrated action editor constructs action sequences by combining them after consulting with the executive producer or the action director. Until the producer or director is completely satisfied, the operator has trial and error of constructing action sequences. In this phase, the 3D-space of the actual location site are also constructed using the VR-space layout tool. The action sequences are arranged in the constructed VR space to confirm spatial relations between performing characters and environments

[Phase 3: Shooting of MR-PreViz]
Before shooting MR-PreViz movies, the operator of 3D-layout tool must fine adjust the constructed virtual spaces to the real structures of shooting site. MR-PreViz movies are shot by the image capturing and compositing camera system. Next, camera-work information is recorded as CWML data. The movie crew including the director and the camera man can test various patterns of camera-work and camera blocking. By viewing the SD level MR-PreViz movies, the director and a cameraman can discuss thr camera-work extemporarily.

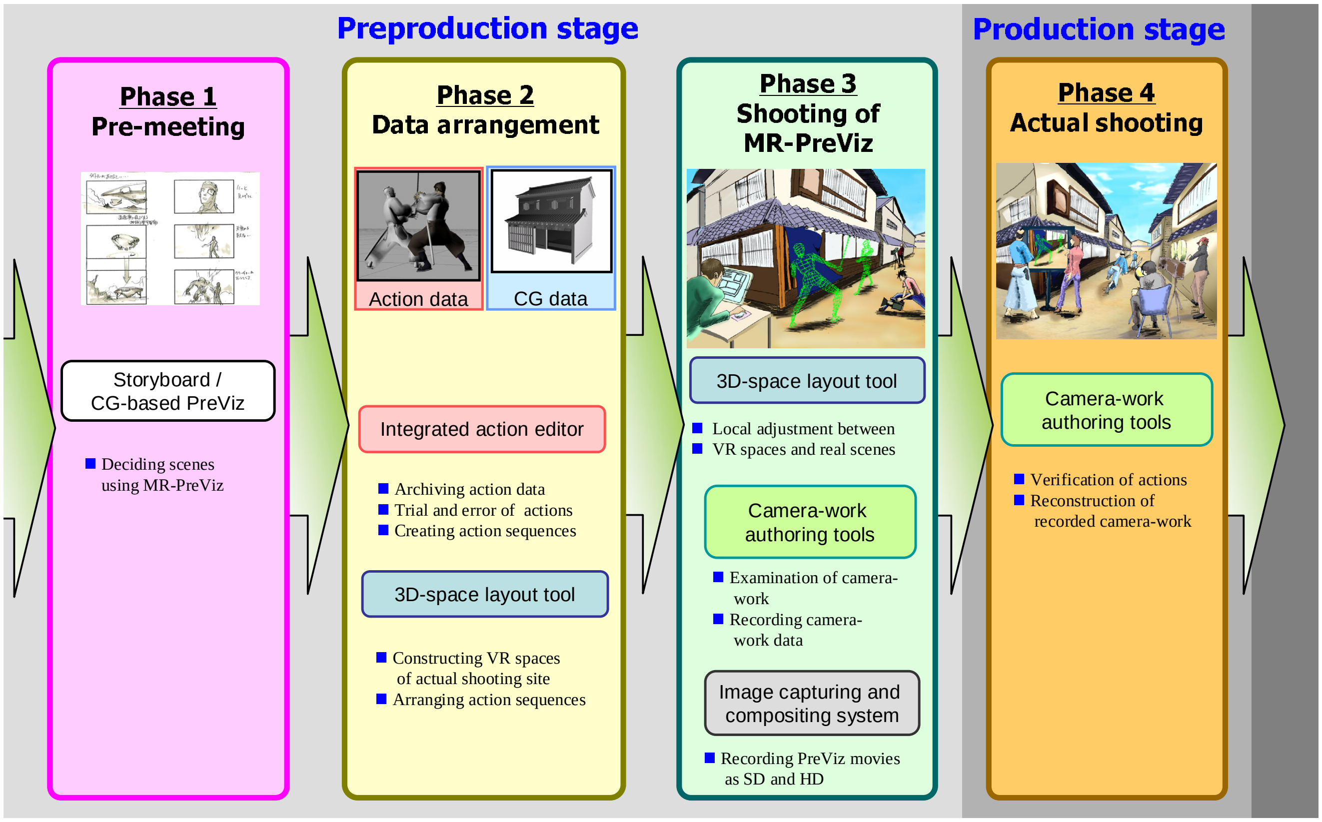


Figure 10: Workflow of filmmaking using MR-PreViz

(a) (b) (c)

Figure 11: Producing MR-PreViz movie in open set

(a) (b) (c)

Figure 12: Composed images

[Phase 4: Actual shooting]

As shooting actual films, preliminarily recorded camera-work and camera blocking are reconstructed using the MRP browser. At the shooting site, MR-PreViz movies are shown to actors to verify their performances immediately before the actual shootings. According to us, our MR-PreViz allows us to save a lot of actual shooting time.

5 TRIAL SHOOTING OF MR-PREVIZ MOVIE

We had a shooting of MR-PreViz movie at the Toei Kyoto studio park. Toei Kyoto studio park is a tourist spot of reconstructed ancient Japanese streets. This park is sometimes used for actual shooting samurai movies or dramas. We merged a 3D video of two samurais' swordfight with real scene images of the park. For this shooting, we used the camera unit and the SD level image compositing unit. Figure 11 shows scenes shot at this location. These are the assumed scenes where in the director and the cameraman discuss the camera-work and camera blocking. The system operator is in front of the monitor and operates the system following the instructions of the director and the cameraman. Figure 12 shows screen shots of MR-PreViz with different camera positions.

Through this trial shoot, we rediscovered issues faced to us. Here, we used only SD level image capturing and compositing system with pan/tilt encoder. The alignments of real and virtual coordinates were operated manually. Our most important issue is researching quick registration method, especially, which can be used for outdoor shooting site.

6 CONCLUSIONS AND FUTURE WORK

The main reason for presenting the outline of our project at an early stage, is that we believe such a new application can provide a new horizon for the AR/MR technology. Filmmaking is a very attractive but an extremely demanding career. Thus, various challenging problems may be encountered while proceeding with this project. By revealing such problems, we can expect the AR/MR research community to provide a solution. In such a case, the foundation of the AR/MR technology would further strengthened. Some such problems are:

- It is not advisable to rely on mechanical sensors when increasing the degrees of freedom of the camera-work from two to three and four. The vision-based tracking method is considered promising while considering the degree of freedom in outdoor environment.

- If we can attain geometric correspondence by the vision based tracking method [12] or some other method, a method to achieve photometric correspondence between outdoor scenes and CGIs should be devised. Even if our MR-PreViz movies are successful, the weather conditions may differ from those of the actual shoot.

- Our MR-PreViz camera system can detect the zooming and focusing factor in real time. These parameters can be applied to CG characters to express some shooting techniques using defocusing. Furthermore, we will improve our system such that it can express lighting techniques used during actual shootings.

- Data acquisition while wearing movie costumes is a big advantage of the 3D Video method. On the contrary, this immense data is a crucial disadvantage. New data compression methods should be developed so that we safely store the necessary footage without degrading its quality.

The PreViz technology cultivated throughout this project can be applied to moving pictures other than films. In addition, theatrical performances or live events in outdoor environment can be effectively pre-visualized and successfully simulated.

ACKNOWLEDGEMENTS

This research is supported in part by Core Research for Evolutional Science and Technology (CREST) Program "Foundation of technology supporting the creation of digital media contents" of Japan Science and Technology Agency (JST). The authors would like to thank Mr. Taneda, Mr. Ogawa, Mr. Beppu, and Mr. Kawano for partly implementation.

REFERENCES

[1] H. Tamura: "Overview and final results of the MR Project, " Proc. 2nd Int. Symp. on Mixed Reality, pp. 87-94, 2001.

[2] H. Tamura, H. Yamamoto, and A. Katayama: "Mixed reality: Future dreams seen at the border between real and virtual worlds, " IEEE Computer Graphics and Applications, Vol. 21, No. 6, pp. 64-70, 2001.

[3] http://www.arvika.de/www/e/home/home.htm

[4] T. Ohshima, T. Kuroki, T. Kobayashi, H. Yamamoto, and H. Tamura: "2001: An MR-Space Odyssey: Applying mixed reality technology to VFX in filmmaking, " SIGGRAPH 2001 Conference Abstracts and Applications, p. 142, 2001.

[5] M. Shin, B. Kim, and J. Park: "AR Storyboard: An augmented reality based interactive storyboard authoring tool, " Proc.4th Int. Symp. on Mixed and Augmented Reality, pp. 198-199, 2005.

[6] S. Feiner, B. MacIntyre and D. Seligmann: "Knowledge-based augmented reality, " Communications of the ACM, Vol. 36, No. 7, pp. 52-62, 1993.

[7] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Juiler and B. MacIntyre: "Recent advances in augmented reality," IEEE Computer Graphics and Applications, Vol. 21, No. 6, pp. 34-47, 2001.

[8] S. Moezzi, L. Tai, and P. Gerard: "Virtual view generation for 3D digital video, " IEEE Multimedia, Vol. 4, No. 1, pp. 18-26, 1997.

[9] T. Kanade, P. Rander, and P.J. Narayanan: "Virtualized reality: Constructing virtual worlds from real scenes, " IEEE Multimedia, Vol. 4, No. 1, pp. 34-47, 1997.

[10] T. Matsuyama and T. Takai: "Generation, visualization, and editing of 3D video, " Proc. 1st Int. Symp. on 3D Data Processing Visualization and Transmission, pp. 234-245, 2002.

[11] R. Ichikari, K. Kawano, A. Kimura, F. Shibata, and H. Tamura: "Mixed reality pre-visualization and camera-work authoring in filmmaking, " Proc. 5th Int. Symp. on Mixed and Augmented Reality, 2006.

[12] M. Oe, T. Sato, and N. Yokoya: "Estimating camera position and posture by using feature landmark database, " Proc. 14th Scandinavian Conf. on Image Analysis (SCIA2005), pp. 171-181, 2005.