The Journal of the Society for Art and Science Vol. x, No. x, pp. xxxx-xxxx

Forensic CG Video Generation with Augmented Reality Technology

Atsushi Sugiura(Non-member) **Masahiro Toyoura**(Non-member) **Xiaoyang Mao**(Member)

Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi
{g12dhl02, mtoyoura, mao} ＠ yamanashi.ac.jp

Abstract

This paper presents a technology for supporting in-house forensic CG video generation by police and prosecution. For confidentiality, it is usually not allowed to generate criminal CG video through subcontracting, whereas existing CG modeling and video editing systems are too complicated to use for police or prosecution. The proposed technique solves this problem with Augmented Reality (AR) technology. We propose to deploy multiple cameras and AR markers on criminal site. While a user of the police or the prosecution plays a role of the criminal wearing a video see-through head mount display, video is automatically generated. CG objects, such as the virtual victim, are displayed on the head mount display and rendered in the video. Our main contributions in this research are concurrent displaying of first-person and third-person views for visualizing spatial relationship between the user and its surroundings, and complementary use of marker information from multiple cameras for solving the occlusion problem inherently caused by operating at the actual criminal site.

Keywords

Augmented Reality, Multi-Marker Complement, Multiple Views, ARToolKit.

1. Introduction

The lay judge system has been introduced in many countries. The lay judges are chosen from citizens in general. They discuss about guilt or innocence and the appropriate punishment with professional judges. The system makes the judge to follow the sense of ordinary people. It also makes the period much shorter. Since the lay judges are amateur, the materials provided to them are required to be easily understandable without expert knowledge.

We focus on the demonstration CG video as a part of forensic materials, which enables the judges to understand the criminal situation intuitively. Although it is already generally used in TV programs, it has been rarely used in the court. This is because subcontracting is inappropriate for confidentiality, while police or prosecution is not expertise for generating movie. We propose to support in-house CG video generation with Augmented Reality (AR) technology. AR technology support to display CG objects in video captured real world. Supplementary information can be added in the video. An initial version of this paper was presented at NICOGRAPH international 2012 [1]. We suppose multiple cameras and multiple markers are deployed at the criminal site. A user, such as the police or the prosecution, plays the role of the criminal wearing a video see-through Head Mounted Display(HMD). CG objects are rendered in the video captured criminal site and displayed to the user with the HMD. In this way, modeling of human motion and actual environment is not necessary, and all things the user should do for generating CG video is playing the role of the criminal on site. AR markers may be occluded by the user and objects in a real site. We solve the problem by complementing marker information among multiple cameras. Displaying multiple views from the multiple cameras enables the user to grasp the spatial relationship between the user and his/her surroundings, even if the user is not well trained for using AR system.

In the next section, we first introduce related works. Section 3 describes how the proposed system generates CG video. The results of subject studies are described in Section 4, and Section 5 concludes the paper.

2. Related works

Several research groups have proposed to use CG models as forensic materials. Section 2.1 introduces the existing work on creating CG models based on the investigation data of criminals.

We use multiple views to help the user understand the spatial relationship between himself/herself and surroundings. The system consists of multiple cameras and multiple PC connected through a client/server network. Existing works on multiple views AR are described in Section 2.2. Research works on reducing network delay for AR applications are introduced in 2.3.

2.1 Crime Scene Modeling

Creating the 3D CG model of criminal site requires technical skills. It also costs much time. March et al. proposed a technique for converting forensic pathology descriptions into 3D CG models [2]. Thali et al. generated individual models of corpses using markers [3]. Captured images, MRI (Magnetic Resonance Imaging) and CT (Computed Tomography) data are integrated to create the 3D CG model with the markers. Compared to the corpses, criminal site and the animation depicting the process of a crime are much more difficult to model since they varies case by case and there are no standard models are available. Se and Jasiobedzki [4] have proposed the instant Scene Modeler (iSM), a 3D imaging system that automatically creates 3D models using an off-the-shelf hand-held stereo camera. The user points the camera at a scene of interest and the system will create a photo-realistic 3D calibrated model automatically within minutes. Clair E et al. [5] proposed to model the crime scenes using the sketch based modeling tool called SketchUp [6]. These tools are good at creating static CG models, but poor at creating dynamical models such as human action or moving weapons. Those dynamical models are vital for depicting a crime.

We solve the problem by using AR technology. In our system, the CG model of victim, which can usually be obtained by modifying a standard model, is rendered into the captured image of crime site, and a video depicting the process of crime can be automatically generated simply by having a user playing the role of criminal. CG objects can be arranged in the site intuitively by using AR markers.

2.2 AR Systems supporting multiple views

Supporting of two or more views in AR system has been actively discussed. Sport is a good application of AR. The 3D soccer observation system developed by Inamoto and Saito generates free viewpoint video from images captured with a large number of cameras [7]. A viewer can specify an arbitrary viewpoint and the system generates the corresponding image by interpolating multiple images from different cameras. Cameras are calibrated in advance. The video generated by interpolating multiple views gives virtual views, which is not suitable for being used as forensic materials since the reliability is most important for the judge. Our system generates third-person views from real cameras. Viewers can change the camera, and examine the situation from the viewpoints of different cameras.

Integrating information of multiple views is another aim of our system. Only from the first person view, the performer is difficult to perceive the depth information. By displaying the third-person views, we expect it would be easy for the performer to keep an intended posture. Thoung et al. [8] proposed an outdoor AR system for investigating the objects at a distance. They adopted zooming cameras in addition to the head mounted cameras. The system can provide the images of zooming cameras. Enomoto et al. [9] employed the cameras of other points of views for diminished reality. By integrating the images from multiple viewpoints, an object behind a marker can be transparently displayed in the first-person view. In our proposed system, we expect the performer to integrate multiple views in real time. The system just provides multiple views, avoiding lowering the performance with the extra cost of multiple view integration. For achieving such purpose, the position and size of the images on HMD should be carefully designed and real time transfer of those images is also very important.

2.3 Reducing Network delay for AR applications

Research works combining AR and network technologies have also been conducted. The work by Okamoto et al. [10] transmits and receives the information on the real world between remote sites and aimed at supporting all operations in real-time. However, they did not address the problem of video latency caused by network transfer, but focused on the latency mainly related to AR applications. Lai and Duh [11] have reported that the latency affects dynamic quantitative perception of virtual objects. The delay problem has been well dealt in mobile AR applications [12]. Pasman and Jansen [13] proposed coarse-to-fine representation of virtual objects for AR applications. In their application, 3D CG models were rendered in a mobile device. Mito et al. [14] transferred minified and compressed images in JPEG format to reduce the network latency in their application of immersive airship controlling. However, they have not examined the effect of network latency for AR airship controlling.

Our system also performs the data communications between personal computers using a network, considering the requirement of synchronized displaying of third-person view images and first-person view images, the size of transferred images should be small. We set the size of transferred image as 160×120 pixels, which is 1/16 of the original captured size. We examine the effect of size reduction in Section 4.4.

3. Proposed method

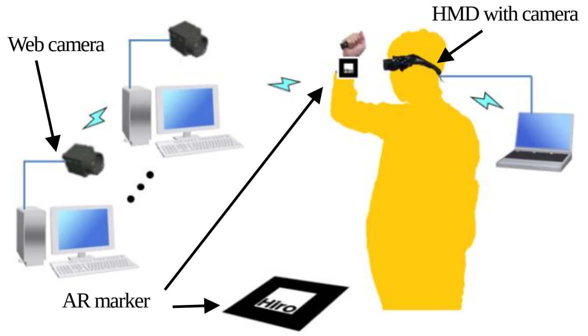
Our method generates a video for visualizing a crime scene with AR technology. The video consists of the images obtained by rendering the virtual objects, such as the 3D model of a victim, weapon, over the real camera captured images of criminal site. The knowledge about CG modeling is not needed.

When using the actual criminal site as the background, we cannot expect large equipment such as position sensors or a good lighting condition. We should detect AR markers in such situation. The problem is solved by using multiple markers captured by multiple cameras in a complementary way. The multiple cameras also support a user to act in the space. The user wears a HMD, and the images from multiple cameras are displayed on the monitor of HMD.

Section 3.1 describes the overview of our system, and Section 3.2 describes the complement of information among multiple markers. Section 3.3 is about how to support the user by displaying multiple views, and Section 3.4 introduces the configuration of server/client network for transferring multiple views.

3.1 Overview

We arrange multiple cameras and multiple AR markers on the actual crime site. We assume the lightings and the position sensors are not available since the space is not a managed studio. The markers are often occluded by the user or other objects in real world. When multiple cameras are placed in the space, we can expect some of them can detect the marker. Then, marker information can be complemented by integrating the information from those cameras. In addition, the multiple cameras provide multiple views supporting the user to grasp the spatial relationship among the virtual objects and real objects including the user himself/herself.



(a) Before overlaid. The user wears HMD. The multiple cameras and markers are deployed in the space.

(b) After overlaid. A video captured with multiple cameras. The virtual victim and the virtual weapon are displayed on the positions of the markers in the video.

Figure 1. Configuration of the proposed system.

The configuration of our proposed system is shown in Figure 1. As shown in Figure 1(a), the user wears a HMD. The video of first-person view is displayed on the HMD. Multiple cameras which are connected to the network are arranged in the space. They capture the video from the third-person viewpoints. The video is transferred to the user's HMD. A composite video with both of the first-person view and the third-person views is provided to the user.

Multiple markers are placed in the space. The individual markers correspond to individual virtual objects, with the one for a virtual victim set on the floor, and the one for a virtual weapon grabbed in the user's hand. The virtual victim and the virtual weapon are displayed on the positions of the markers in the video. It means that the video captured in the actual crime site is augmented with the virtual objects by¥ AR technology.

3.2 Complement of multiple marker information

ARToolKit [15] is a library for creating AR environment easily. It detects a marker with a camera, estimates the position and posture of the marker, and displays a virtual object on the position indicated by the marker. However, when a part of the marker is absent from the field of view of the camera or when the mark is in an inappropriate angle relative to the camera, the marker becomes undetectable. We solve the problem by using the information from multiple cameras complementary.

Figure 2 shows the flow of the complement of marker information.

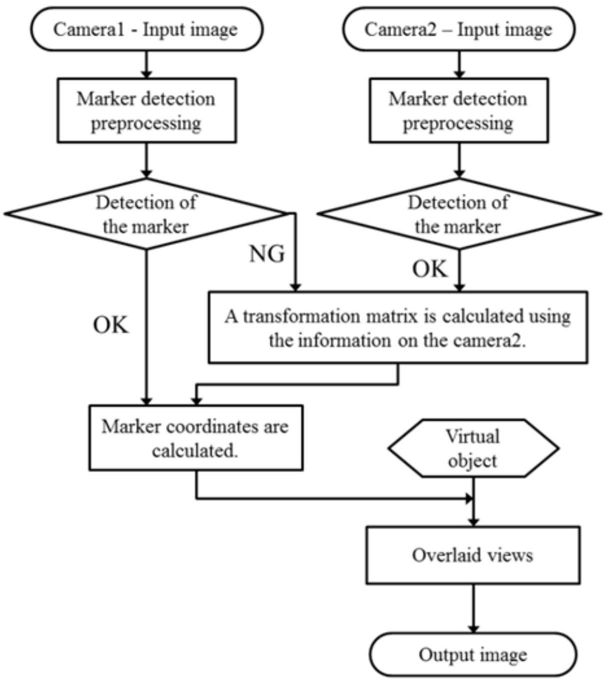


Figure 2. Complement of undetectable marker information. When camera 1 fails to detect a mark, a geometric transformation matrix is calculated using the information from camera 2.

When a marker is detectable, a geometric transformation matrix P can be calculated with the function of ARToolKit. Then a virtual object can be displayed on the marker in the marker coordinate system.

Without loss of generality, let's consider a system consists of two cameras and two markers as shown in Figure 3. Denote the geometric transformation matrix of marker A and B detected from camera 1 are P_{1A} and P_{1B}, the geometric transformation matrix of marker A and B detected from camera 2 are P_{2A} and P_{2B}. Assume camera 1 can detect both marker A and B, but camera 2 can detect marker B only. Because marker A is undetectable from camera 2, geometric transformation matrix P_{2A} cannot be calculated. Assume the camera coordinate system of camera 1 viewed from camera 2 is $\Phi_{2\to1}$, then the transformation matrix of marker B viewed from camera 2 via camera 1 becomes $\Phi_{2\to1} \cdot P_{1B}$. That is, we can obtain the following equation:

$$P_{2B} = \Phi_{2\to1} \cdot P_{1B}. \tag{1}$$

Similarly, the transformation matrix for maker A viewed from camera 2 via camera 1 is

$$P_{2A} = \Phi_{2\to1} \cdot P_{1A}. \tag{2}$$

Combining equation (1) and (2) we obtain the following equation:

$$P_{2B} = \Phi_{2\to1} \cdot P_{1B}. \tag{3}$$

Since P_{1A}, P_{1B}, and P_{2B} are available, P_{2A} can be computed with equation (3) and the marker coordinate system of marker A viewed from camera 2 can be calculated from P_{2A}.

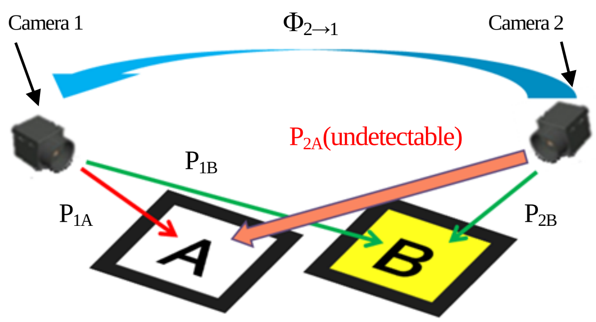


Figure 3. Marker information complement. The projection metrics by the two cameras and two markers can be calculated by the complement of marker information.

The principle behind the above calculation is that we can first obtain the spatial relationship between the two cameras through the marker which is detectable by both cameras (Equation (1)) and then compute the coordinated system of the occluded marker from its coordinate viewed from the other camera. Therefore, to detect an occluded marker, at least one more marker should not be occluded from both cameras. Moreover, there is no way to complement a marker which is occluded from all cameras.

In the proposed system, a network of client/server model is constructed for communicating among cameras. Marker information is detected by the client and sent to a server. A server receives data from a client, when there is a marker which was undetectable, the server complements marker information by using the data from other clients.

3.3 Displaying multiple views to the user

When creating a synthetic image, the spatial relationship between a real object and a virtual object becomes important. It may be difficult for a performer to precisely grasp the spatial relationship only with the image of a first person viewpoint. We improve the situation by providing the performer the image of a third person viewpoint. Since the performer can see the image of various viewpoints, it becomes easier for him/her to understand the spatial relationship among virtual objects and actual criminal site. The multiple views are also useful for helping the user to confirm whether the numbers and positions of cameras and markers are appropriate during the setup of the environment, which is usually a difficult task especially when the actual crime site has a complicated structure.

3.4 Network Data transfer between cameras

A network is used in order to transmit images of two or more viewpoint between personal computers. A server/client model using TCP socket connection is used. The image of the third person view acquired with the camera of a client computer is sent to a server via network. If the size of the images being transferred is large, the system can become slow depending on the traffic of network. However, real time processing is very important in case of an AR system. We solve the problem by using small images for the third-person view. The downsizing is executed before the images are transferred.

4. Experiments

We examined whether the third person view can help subjects grasping the spatial relationship between the subjects and virtual objects. Complement of marker information between multiple cameras and multiple markers is also examined through another subject study.

4.1 Environment

Experimental environment is shown in Figure 4. Three PCs are deployed in a room. A HMD with 2 cameras (Wrap920AR, Vuzix Corporation) is connected to a server PC. The resolution of cameras is 640×480 and frame rate is 30fps. The resolution of display monitors is 800×600. The HMD provides the video of the first-person view to the subject. The remaining two PCs are clients with web cameras (Logicool C920 and QCam-200SX) connected. They capture the third-person views. The video were 640×480 in resolution and 30fps in frame rate.

Two markers are placed in the space, one of them represents the position of a virtual victim, and the other the position of a virtual weapon. Figure 5(a) shows the marker for the virtual victim, and Figure 5(b) the marker for the virtual weapon. The size of both markers is 80mm×80mm. If the marker for the virtual victim is placed on the floor, it rarely came into the sight of the subject. Therefore, the marker was placed on a stand as shown in Figure 5(a). We designed the weapon as shown in Figure 5(b). The edge of the weapon was rendered as a virtual object.

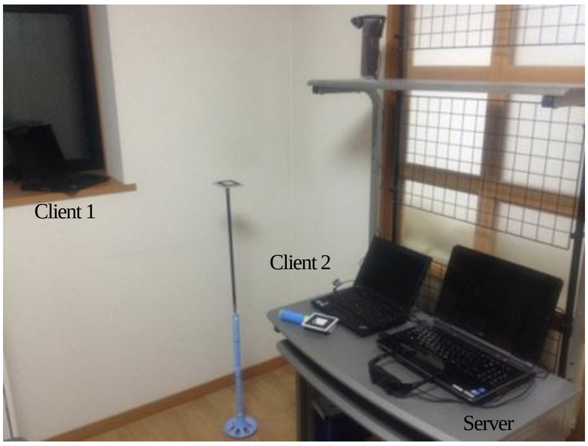
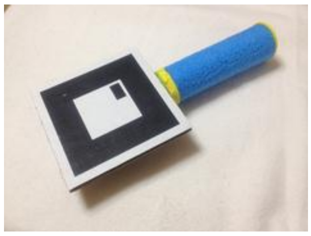


Figure 4. Environment for experiments. A server PC with HMD and two clients PCs with web camera are deployed in a room. The marker information is acquired from the cameras.

(a) Victim marker



(b) Weapon marker

Figure 5. Markers used in experiments. The size of both markers is 80mm×80mm. The height of the victim marker is 1.1m. A handle is attached to the weapon marker.

4.2 Complement of marker information

In this experiment, we examined whether the complement of marker information based on the server/client communication was successful.

An operation scene is shown in Figure 6. The user wore a HMD and grabbed the weapon marker in his hand. The cameras on the HMD captured the video of first-person view. The video of the third-person view from other cameras were transferred to the server PC.

The victim marker was arranged in the detectable space from all cameras. In several frames, the weapon marker was out of sight in first-person view. The result of marker complement is shown in Figure 7.

Figure 6. Operation scene. The user wore a HMD and grabbed the weapon marker in his hand. Three cameras capture the action.

(a) Result without the complement of marker information. Since the weapon marker went out of sight of HMD cameras, the virtual weapon could not be displayed.

(b) Result with the complement of marker information. The information of undetected weapon marker was complemented by integrating marker information detected in other cameras. The virtual weapon could be displayed even if the marker went out of the view of HMD cameras.

Figure 7. Marker information complement.

(a) First-person view.

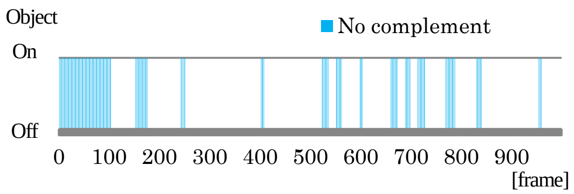
(b) Third-person view.

Figure 8. Complement of marker information based on the victim marker. The marker information is calculated using the victim marker of both first-person view and third-person view.

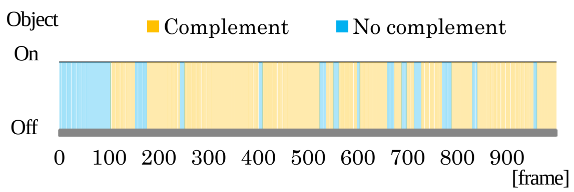
Figure 7(a) shows the result without marker complement. Since the weapon marker went out of sight of HMD cameras, the virtual weapon could not be displayed. Figure 7(b) shows the result with marker complement. The virtual weapon could be displayed even if the marker went out of the view. The marker information from the client computer cameras was transferred to the server PC. The information of undetected weapon marker was complemented by integrating marker information detected in other cameras.

Figure 8(a) represents the first-person view, and Figure 8(b) the third-person view. Based on the victim marker, which could be detected in both two cameras, the weapon marker is displayed in its position although the marker could not be detected in the camera of the first-person view. The information about the spatial relationship between the victim marker and the weapon marker is transferred through the network. As the result, the virtual weapon could be displayed in the first-person view.

We examined the scene that the subject stabbed the virtual victim with the virtual weapon. Whether the virtual weapon was detected or not is shown in Figure 9. The marker could not be detected in 765 frames without complement of marker information (Figure 9(a)), but could be detected in all frames with the complement of marker information (Figure 9(b)). We conclude that the complement was effective for displaying the marker for CG video generation in our proposed system.



(a) The result without the complement of marker information.



(b) The result with the complement of marker information.

Figure 9. The result of marker detection with/without complement of marker information.

4.3 User support by displaying multiple views

In this experiment, we examined whether displaying multiple views is effective for helping user grasping spatial relationship between the virtual and real objects.

2 male subjects and 1 female subject in their 30's joined the experiment. Figure 10 shows the image shown on HMD. The third-person views are not displayed in Figure 10(a), but displayed in Figure 10(b).

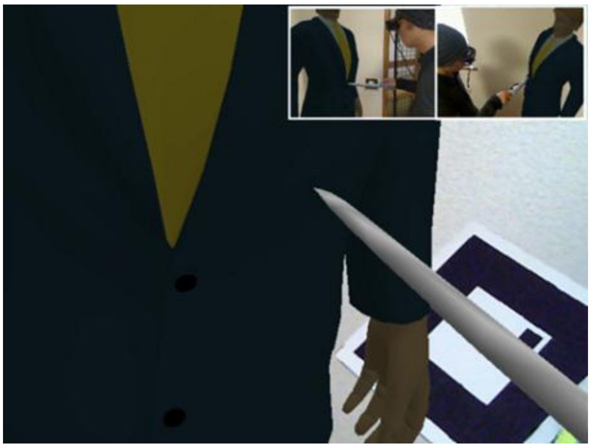
From the comments of all subjects, displaying third-person views helped to easily grasping the spatial relationship between the virtual objects and real objects including the subjects themselves. As the task we had the subjects touch a specified point with the tip of the virtual weapon. A 250mm×250mm×250 mm cube was displayed on the victim marker. The first task was to put the tip of the virtual weapon on the surface of the cube. The second task was to put the tip at 50mm distance away from the surface of the cube. We examined the accuracy and time to complete the tasks with and without third-person views displayed. When the virtual objects are 300mm away from each other, we start to measure the time for the tasks. We stop timing when the subjects stopped the motion. The scene of the tasks is shown in Figure 12.

Results are shown in Figure 13 and Figure 14. The error of distance between virtual objects became smaller by displaying the third-person views. The time for the tasks also became shorter. Visualization of the third-person views was confirmed to be effective for the tasks.

To furthermore examine the relationship between the difficulty of tasks and the performance of the user, we set another target angles of stabbing. We had the subjects stab the virtual cube in each of the following 4 angles: 30 degree right, 45 degree right, 30 degree up and 45 degree up. The error and time were measured at the time that the virtual weapon touched the cube. The average error and time of the tasks are summarized in Table 1. According to Table 1, third-person views could help the subjects to set the virtual weapon at more correct angle in shorter time.

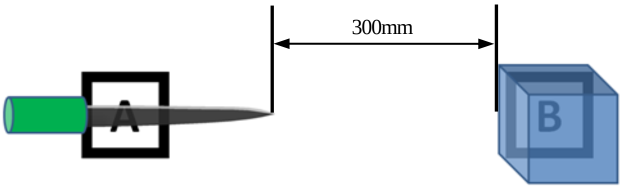
The size of the cube was also changed between 100mm and 250 mm on a side to confirm the relationship between the size and the performance of subjects. The results are shown in Table 2. In both cases, the subjects could set the virtual weapon at more correct angle in shorter time. The larger targets tend to be stabbed in shorter time, but the accuracy of the position had no big difference relative to the size of the target.

(a) Without third-person view. Grasping spatial relationship is difficult with first-person view only.

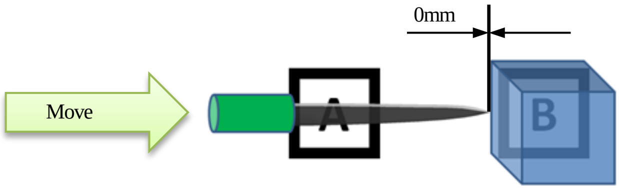


(b) With third-person views. The third-person views are added to the upper right of the image displayed to the HMD. Grasping spatial relationship become easier by referring to the third -person views.

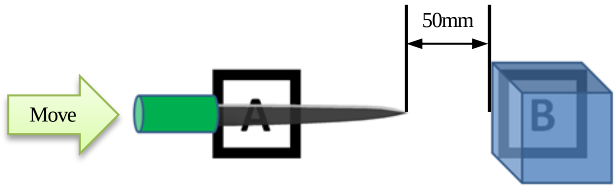
Figure 10. Images displayed to the HMD.



(a) Before Task

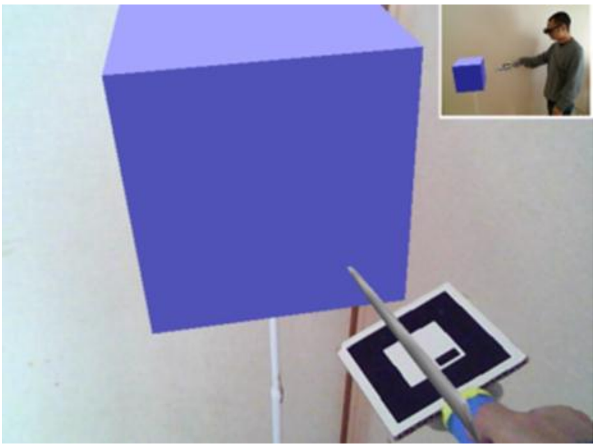


(b) Task 1

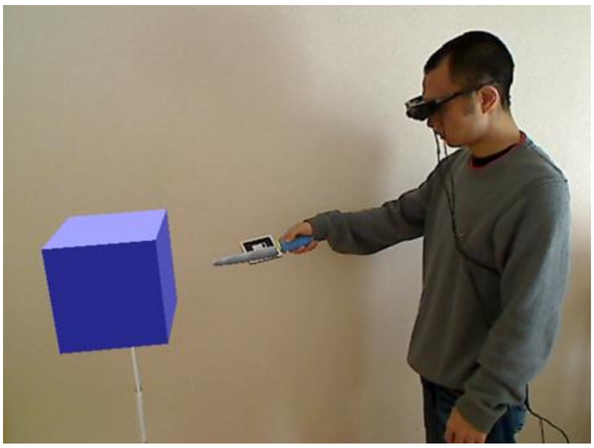


(c) Task 2

Figure 11. Tasks for evaluating the effectiveness of displaying third-person views.



(a) With third-person view.



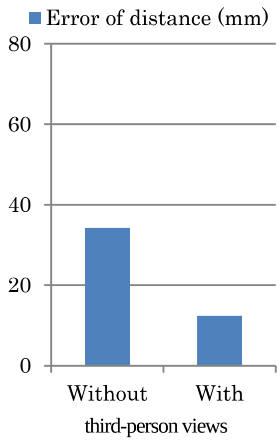
(b) An example of third-person view.

Figure 12. Experiment scene. We evaluated whether the subjects could grasp the spatial relationship. The virtual weapon was brought close to a cube. The time and distance to the target were measured.

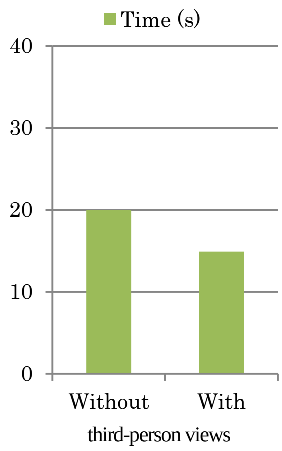
A problem found in experiments was that a planer weapon marker as shown in Figure 15(a) may not be detected depending on its orientation. As shown in Figure 15(b), we solve the problem by arranging five markers on a cube. As the result, it can be robustly detected from any direction and the subject could play a role of the criminal without paying special attention to the marker orientation.

We also considered the position of the third-person views in the display monitor. Lower right area often overlapped with the right hand of the users. Therefore, for the right-handed users, we use the upper right area of the monitor.

The relative distance and angle between the weapon and victim are also displayed on the monitor as shown in Figure 16. In real application, a user can use this function for checking the consistency of current positions to those of an autopsy report. The distance and angle are calculated from the positions of their corresponding markers. A user can select whether to display the information on the monitor or not.

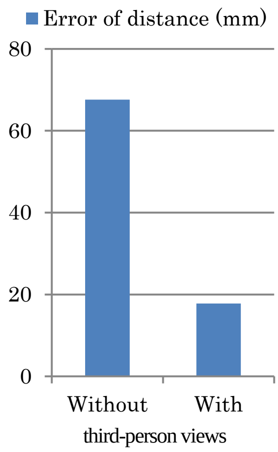


(a) Error of distance

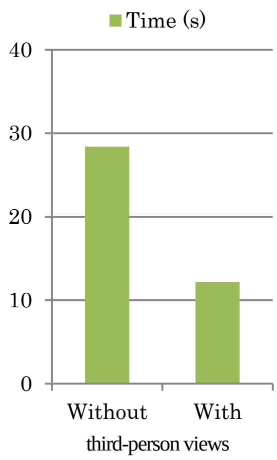


(b) Time

Figure 13. Result of task1.



(a) Error of distance



(b) Time

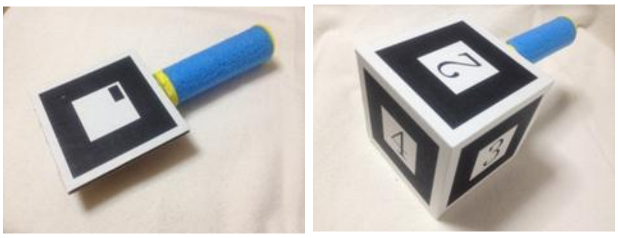
Figure 14. Result of task2.

Table 1. Result for different stabbing angles.

	Error of Angle (°)		Time (s)	
Third-person views	Without	With	Without	With
30° right	4.82	2.79	16.24	15.27
45° right	6.48	1.97	18.04	15.4
30° up	4.78	2.23	16.82	15.38
45° up	6.14	2.08	17.9	16.12

Table 2. Result by varying target size.

	Error of Dis (mm)		Time (s)	
Third-person views	Without	With	Without	With
100mm	31.8	9.67	33.91	24.9
250mm	33.96	12.37	19.94	14.87



(a) A single planar marker (b) Markers on a cube

Figure 15. Marker for the virtual weapon. A single planar marker often get missed from third-person view cameras. By arranging 5 markers on a cube, the subject could play a role of the criminal without special attention to the marker direction.

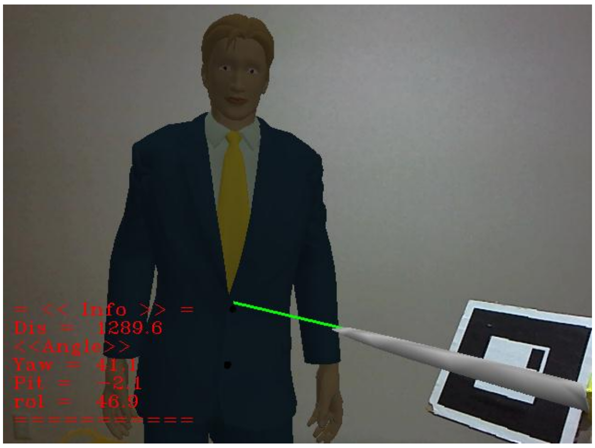


Figure 16. Distance and angle information displayed. The user can check the consistency of current positions to those given by an autopsy report.

4.4 Network delay and image size

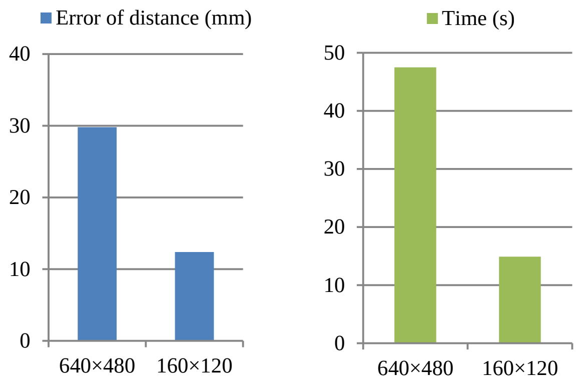
We propose to transfer third-person view images in smaller size for avoiding network delay. Since the images are transferred through general LAN, it takes longer time when images of larger size are transferred. The network delay may make it difficult for the user to positioning virtual weapons accurately.

We first measured the time for transferring the images of 640×480, 320×240 and 160×120, respectively. Table 3 shows the result. The time was measured as the average of 10 images. They include compression time in a client computer, which is almost linear to the number of pixels, and was ignorable.

Table 3. Delay time when images of 640×480, 320×240 and 160×120 are transferred through our network.

Image size	delay time (s)
640×480	0.64
320×240	0.28
160×120	0.08

Next, we had the subjects perform the stabbing task of the experiment described in Section 4.3. The accuracy and time for 640×480 and 160×120 are shown in Figure 17. The larger the transferred images are the longer the task completion time is. This is because the network latency makes it difficult for the subjects to determine the position of the virtual weapon relative to the virtual objects. Figure 17 shows that the subjects could set the virtual weapon at more correct position in case 160×120 images were transferred and displayed. This demonstrates that transferring downsized images is effective and sufficient for supporting accurate operations.



(a) Error of distance (b) Time

Figure 17. Comparison of accuracy and completion time for different images sizes.

4.5 Remarks by police

We discussed with local police about the potential of using our system in real situation. To the best of their knowledge, currently no such techniques are available and Judges are preceded mainly with descriptions and photographs. They think an AR system is particularly useful for stabbing cases since it is impossible to reenact the injury inside the victim's body. On the other hand, they also pointed that functions for guaranteeing the consistency between the generated video and the fact should be provided. Taking the stabbing cases as the example, the angle, speed and depth of stabbing are very important information for validating the evidence, and should be visualized in consistency with autopsy report or other evidences. Improving the system based on those remarks is one of our major future research directions.

5. Conclusions

In this paper, in-house 3D CG video generation supporting system using AR technology is proposed. The system focuses on the reenactment of the crimes causing injury or death, which are impossible to reenact actually. PCs connected with a server/client network provide both of the first-person view and the third-person views on the HMD. It supports users to grasp the spatial relationship between themselves and surroundings. The marker information detected from multiple markers is integrated in the server PC. When a marker is occluded from a camera, its information can be obtained from another camera.

As a limitation of our current system, the victim does not react to the user. Actual victim may escape or be stunned, go down when the weapon is approaching. The victim of our current system just stands up on the marker. We can replace the victim with captured data in advance. Timing of reaction can be defined by the approaching distance and angle of the user. Modeling the reaction of the victim would improve the reality of the generated video.

As a future work, the interface for augmenting additional information to the video is desired. It should support the visualization of information in the video. Recently, CT and MRI have been introduced to judicial autopsy. Such data will facilitate the generation of precise 3D models. We plan to provide the technologies for calibrating CG models to real objects.

As another potential application, the lay judges can not only watch the video, but experience the criminal site themselves by using the AR system themselves, so as to understand the situation of the crime better.

References

[1] A. Sugiura, M. Toyoura, and X. Mao, "Forensic CG Video Generation with Augmented Reality Technology," NICO-GRAPH International, pp.94-101, 2012.

[2] J. March, D. Schofield, M. Evison, and N. Woodford, "Three-Dimensional Computer Visualization of Forensic Pathology Data," The American Journal of Forensic Medicine and Pathology, vol.25, pp.60-70, 2004.

[3] M. Thali, M. Braun, U. Buck, and E. Aghayev, "Scientific Documentation, Reconstruction and Animation in Forensic: Individual and Real 3D Data Based Geo-Metric Approach Including Optical Body/Object Surface and Radiological CT/MRI Scanning," Journal of Forensic Sciences, vol.50, no.2, 2005.

[4] S. Se, P. Jasiobedzki, "Instant Scene Modeler for Crime Scene Reconstruction," VCPR, vol.3, pp.123-131, 2005.

[5] E. S. Clair, A. Maloney and A. Schade, "An Introduction to Building 3D Crime Scene Models Using SketchUp," Journal on the Association for Crime Scene Reconstruction, vol.18, no.3, pp.29-47, 2012.

[6] Trimble, SketchUp, http://www.sketchup.com/.

[7] N. Inamoto and H. Saito, "Fly-Through Observation System for 3D Soccer Movie Based on Viewpoint Interpolation," Journal of The Institute of Image Information and Television Engineers, vol.58, no.4, pp.529-539, 2004.

[8] T. N. Hoang and B. H. Thomas , "An Investigation of Camera Position, Visualization, and the Effects of Sensor Errors and Head Movement," ICAT, 2011.

[9] A. Enomoto and H. Saito, "Diminished Reality using Multiple Handheld Camera," ACCV, pp.130-135, 2007.

[10] Y. Okamoto, I. Kitahara, Y. Ohta, "3D Human Body Display by Using Depth Information for Remote Shared Mixed Reality," IEICE Technical Report, vol.94, pp.830-838, 2011.

[11] W.Y. Lai, H.B.L. Duh, "Effects of frame rate for visualization of dynamic quantitative information in a head-mounted display," SMC, vol.7, pp.6485-6490, 2004.

[12] G. Papagiannakis, G. Singh, N.M. Thalmann, "A survey of mobile and wireless technologies for augmented reality systems," Computer Animation and Virtual Worlds, vol.19, no.1, pp.3-22, 2008.

[13] W. Pasman, F. W. Jansen, "Distributed Low-Latency Rendering for Mobile AR," IMAR, 2001.

[14] H. Mito, K. Yamazawa and N. Yokoya, "Support System for Controlling an Unmanned Airship Using Visual Information Based on Mixed Reality," IEICE TechnicalReport, vol.108, no.432, pp.163-168, 2009.

[15] H. Kato and M. Billinghurst, "Marker tracking nad hmd calibration for a video-based augmented reality conferencing system," International Workshop on Augmented Reality, pp.85-94, 1999.

Atsushi SUGIURA received the B.S. and M.S. degrees in engineering from University of Yamanashi, Japan in 2002 and 2012, respectively. He is a Ph.D. student in Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Japan. His research interests include augmented reality and forensic visualization.

Masahiro TOYOURA received the B.S. degree in Engineering, M.S. and Ph.D. degrees in Informatics from Kyoto University in 2003, 2005 and 2008, respectively. He is currently an assistant professor at Interdisciplinary Graduate School of Medical and Engineering, University of Yamanashi, Japan. His research interests are augmented reality and computer/human vision. He is a member of IEEE Computer Society.

Xiaoyang MAO received her B.S. in Computer Science from Fudan University China, M.S. and Ph.D. in Computer Science from University of Tokyo. She is currently a professor at Interdisciplinary Graduate School of Medical and Engineering, University of Yamanashi, Japan. Her research interests include texture synthesis, non-photo-realistic rendering and their application to scientific visualization. She is a member of ACM SIGRRAPH and IEEE Computer Society.