

Dynamic Projection Mapping on Multiple Non-Rigid Moving Objects for Stage Performance Applications

Ryohei Nakatsu¹, Ningfeng Yang¹, Hirokazu Takata², Takashi Nakanishi², Makoto Kitaguchi³, and Naoko Tosa⁴

¹ NT & Associates, Kyoto, Japan

² TakumiVision, Kyoto, Japan

³ TECHMAC, Kyoto, Japan

⁴ Graduate School of Advanced Integrated Studies in Human Survivability, Kyoto University, Kyoto, Japan

nakatsu.ryohei@gmail.com, yangnf@gmail.com,
takata@takumivision.co.jp, nakanishi@takumivision.co.jp,
makoto.kitaguchi@techne-magic.co.jp, tosa@media.kyoto-u.ac.jp

Abstract. In this paper we have proposed and developed a projection mapping system that can project 3D images on multiple moving objects that change their positions and shapes. Although projection mapping has become a new type of image projection method, so far most of the projection mappings have focused on the projection of 3D images on objects such as buildings that stay static without changing their shapes. If projection mapping on non-rigid moving objects such as human performers is realized, application area of projection mapping will become far wider.

To demonstrate this, a simple test of integrating projection mapping and Noh performance was carried out. Although Noh performance has been considered difficult to understand for beginners and foreigners because of its minimalism, by projecting 2D images that suggest inner emotion of a main character and also ongoing story, the evaluation result show that this new type of Noh performance could make it easier to be understood. This means that the integration of projection mapping and various stage performances could open-up new possibilities for various kinds of performing arts to add their values.

To realize this, we have proposed a projection mapping system consisting of multiple depth sensors and multiple projectors. In the first step, using multiple depth sensors the position and shape of multiple objects are detected in real time and the obtained 3D models of the multiple objects are sent to the projection phase. In the projection phase, the images to be projected on the objects are rendered on the obtained 3D model of the object. Then using multiple projectors real-time 3D projection mapping is carried out on the multiple moving objects. The system is now under development and demonstration of actual Noh performance using this technology will be achieved.

Keywords: Projection mapping, Depth sensor, Moving object, Performing arts, Noh performance, Japanese culture, Inner emotion, Media Art.

1 Introduction

Projection mapping is a technology to project 3D images/videos on 3D objects such as buildings [1][2]. As this is one area of Augmented Reality, sometimes projection mapping has been called Spatial Augmented Reality [3]. Conventional projection has been done onto a rectangular flat screen. Once a projection onto a 3D object such as a building is realized by adopting the projection to the shape of the object, it is possible to give the impression that the building has changed into something different.

So far, it seems that the aims of projection mapping have been limited only to entertainment such as a show at a theme park [4] or ceremonial event [5]. In most of these cases images used for the event are made of computer graphics and include entertaining contents. If art contents would be used for projection mapping, however, such projection mapping would create new possibility by changing art exhibition spaces from closed space such as museums and galleries to open public spaces such as city centers, parks, etc. Two of the authors of this paper, Naoko Tosa and Ryohei Nakatsu, have noticed the possibility of projection mapping for such new applications, and have been working on the projection mapping events using Naoko Tosa's media artworks as the contents. Below are two of such events carried out by them.

(1) Projection mapping at ArtScience Museum in Singapore

Naoko Tosa has been working on the creation of media art based on the concept of integrating Japanese tradition and her modern media art, and based on this she has been working on various projects of letting Japanese traditional/modern art to be understood by people not only in Japan but also in overseas. As one of such activities she has created a new media art called "Sound of Ikebana: Four Seasons," in which she let liquid such as color paint jumping up by giving it sound vibration, and then by shooting the generated form using high-speed camera she succeeded in merging her Japanese sensitivity with technologies [6][7]. Funded by METI (Ministry of Economy, Trade and Industry), the projection mapping using this media art as the content was carried out in 2014 at ArtScience Museum in Singapore (Fig. 1). As in Singapore there is no clear difference among each season, this projection mapping focusing the clear transition from one season to another in Japan gave a strong impression about Japan and Japanese culture to the people in Singapore and nearby countries. The event gathered lots of audience and various media in Singapore treated this event.



Fig. 1. A scene of the projection mapping of "Sound of Ikebana" carried out at ArtScience Museum in Singapore.

(2) Projection mapping at Kyoto National Museum

Year 2015 was the 400 anniversary of famous Japanese art school called RIMPA [8] and in Kyoto there were many events celebrating the anniversary. As one of those events Kyoto Prefectural Government called for a projection mapping event. Submitting our proposal responding to this call for event and being accepted by the local government, Naoko Tosa and Ryohei Nakatsu have created a media art called “Thunder God and Wind God in 21st Century.” Using this media art as the content, they carried out the projection mapping in March 2015 at Kyoto National Museum [9].

When a RIMPA exhibition is carried out, the usual way is to exhibit one or several of famous RIMPA paintings as the core of the event. However, RIMPA is not a simple heritage that existed long time ago. On the other hand, RIMPA is an art concept that has to be handed to the present and even to the future [8]. Therefore, they tried to create new RIMPA-style content. Also to appeal the modernity of RIMPA, it is adequate to connect technology and content. Therefore they introduced several cutting-edge technologies in the content creation process. For the projection, twenty most recent high-end projectors with more than 20,000 lumen were used to project the created video content onto the two buildings of Kyoto National Museum. The event gathered almost 20,000 people during the four-day event and was successful. Figure 2 shows a scene of the projection mapping.



Fig. 2. A scene of projection mapping of “21 Centuries Wind God and Thunder God in 21st Century” at Kyoto National Museum

Through these event, we came to the belief that art contents very well fits the new image expression method such as projection mapping. In Kyoto, because of its long-time history and tradition, traditional performing art such as “Noh” is frequently performed. Therefore, we came to an idea of integrating Noh performance and projection mapping using art content. Also, this idea could be applied to other traditional performance such as Kabuki, dance performance, etc. and could open a new possibility for new type of traditional performance and performing arts. Unfortunately, so far most of the projection mapping events have been limited to those where objects for the projection mapping do not move and do not change their shapes. If the above idea is to be realized, it is necessary to develop projection mapping technologies that can realize projection mapping on moving non-rigid object. This idea led us to the development of the new technology described in this paper.

The structure of this paper is the following. In Section 2, a simple experiment where integration of Noh performance and 2D projection mapping is realized is described. Based on the evaluation of this experiment, the basic concept of projection mapping on non-rigid moving object is introduced. In Section 3, Related works on projection mapping is described. In Section 4, overview of the whole system, technologies for tracking 3D objects and extracting 3D model of the objects, and technologies for 3D projection mapping on the moving non-rigid objects using the obtained 3D models are described. And Section 5 concludes the paper.

2 Integration of Noh Performance and Projection Mapping: Projection Mapping on Non-Rigid Moving Objects

As is described in the introduction, we, based on the belief that the integration of projection mapping and art would give a new value to traditional performing art such as Noh, carried out the trial of achieving projection mapping onto Noh performers and Noh background.

“Noh” is globally well known because of its condensed and abstracted body movement and speech, which can be called “minimalism” [10]. Because of this minimalism, however, it is very difficult for foreigners and even for Japanese Noh beginners to understand the story development of Noh play. It is a key issue for Noh, if Japan wants to appeal globally the charm of Noh, to be understood by foreigners and Japanese beginners keeping its sophistication and aesthetics. For this difficult challenge, we think that integration of Noh and projection mapping would be able to create a new style of Noh performance in which Noh would be more easily understood and appreciated. Our idea is to project images onto a Noh performer and background. For example, one idea is to project images that would indicate the inner emotion of the main character acted by the Noh performer. Also, it would be interesting to project images that indicate the story development on the background wall.

Based on such concept we have developed a simple system in which 2D projection is carried out on a Noh performer and background wall. In this case it is necessary for the system to trace the position and posture of the Noh performer and to carry out the projection. Figure 3 illustrates the system setting on the Noh stage and Figure 4 illustrates how the trial of the Noh performance was carried out. Although this was a closed performance carried out at Kyoto University, the main character of the Noh performance was acted by a very famous Noh player in Kyoto. Therefore, the aesthetics the original Noh performance had was maintained and the viewers could judge whether projection of two types of images really increased the value of the original Noh performance or not.

In this trial, Microsoft Kinect was used as a position sensor. Microsoft Kinect has the capability of detecting a position and a posture of a person using infrared depth sensor and we have used this function. Based on this trial it was revealed that, as the major application area of Kinect is gaming, the preciseness of the function of detecting person’s position and posture is not satisfactory. Also we recognized that in artistic performance such as Noh, such precision is the key to add values to such performance. At the same time we understood that the angle range of the Kinect for the human detection is narrow and it is difficult for it to trace the movement of a Noh player who moves around a wide Noh stage.

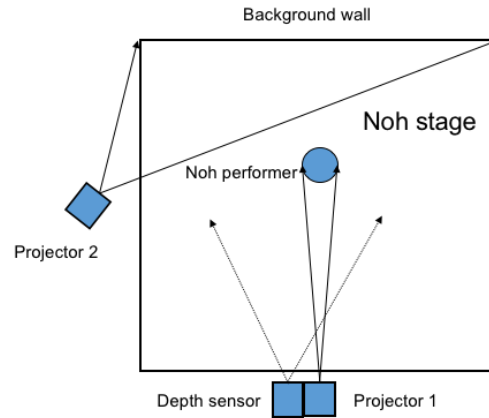


Fig. 3. Stage setting of projection on Noh performance.

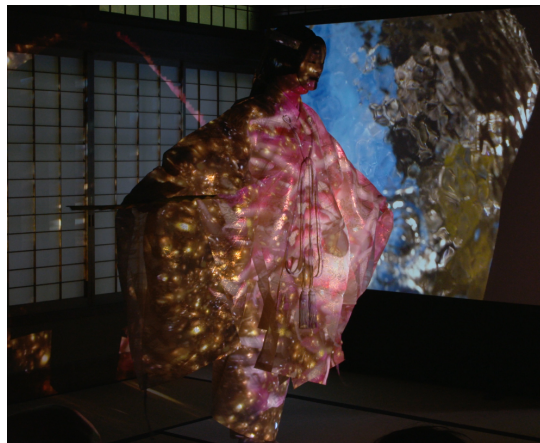


Fig. 4. A scene of trial experiment of projection on Noh performance

At the same time, however, we found that this new trial of integrating projection mapping, although it was a simple 2D projection mapping, and Noh performance was highly appreciated by viewers as a new Noh performing style. Especially international students who joined this experimental trial expressed that, although beforehand they thought Noh is difficult to understand because of slow and subtle movements of Noh performers, this new Noh style was easier for them to understand both the ongoing story and also the inner emotion of a Noh character. Also they expressed that and, this new trial could create new Noh performance style.

This means that, if we could develop a system that can trace the position and posture of a Noh performer who moves around a Noh stage and carry out projection mapping on the moving Noh performer, we could expand the possibility of not only traditional performance such as Noh but also other stage performances including art and entertainment. Also, when we showed this results to the organizer and director of various

events, they told us that this system could be applied not only to stage performances but also to other areas such as fashion shows, live events, wedding ceremonies, etc. and could add new value and charm to these events/shows. Therefore, we believe that the development of dynamic projection mapping on moving non-rigid objects has the possibility of creating new market for projection mapping.

3 Related Works

Shader Lamps [11] is a pioneering research showing the possibility of projection mapping by making it possible to project images on objects with complicated shapes. Since then various types of projection mapping has been proposed, but most of them have been limited to static and/or rigid objects [12][13].

To overcome these restrictions, recently various studies focusing dynamic projection mapping on moving objects have been reported. The key technology for dynamic projection mapping is how to detect/trace the position and shape of moving objects. There are several ways to approach this key issue.

One is to utilize the texture on the object and by using image processing technologies to trace the position and also the shape of the target object [14][15]. Although these texture-based methods have the advantage that they don't need any additional equipment for the detection of the position and shape of the object, these methods have several problems. For example, the tracking performance depends on the texture itself and its density. Also, the lighting, especially the lighting of projection mapping itself, would influence the accuracy of the tracing very much.

Another method to precisely trace the object is the usage of markers. Compared with texture-based methods, the usage of markers can achieve more robust tracking, as the markers could be designed optimally depending on the type of the object, its movement, etc. [16][17]. Probably one of the most notable research among them is by Narita et al. [18] that presents a projection mapping system that is able to dynamically project onto moving objects. They use a 1000fps camera and galvanometer mirrors to track and illuminate the object with a single projector. The resulting tracking is extremely fast and impressive results are achieved. However, this research needs special equipment such as high-speed camera and high-speed projector, and is limited only to experimental environment.

Recently by the emergence of depth camera that, by using infra-red beamer and detector, can detect the distance between the sensor and many points on the target object. Especially by the introduction of Microsoft Kinect that has this capability and that could be purchased by reasonable price, many researcher have used this sensor for the detection of moving object [19][20]. However, Kinect has been developed mainly aiming the gaming situation in which gamers would do physical actions just in front of the sensor, and therefore its effective range including distance to the object and the angle between the object and the sensor is narrowly limited. Because of this, the usage of Kinect has restricted the trial of dynamic projection mapping to only experimental use.

Compared with these previous achievements, the system we propose and are developing has the below originality.

- (1) Most of the previous works have stayed only at research stage. Some used markers that is cumbersome in actual applications. Others used special hardware such as a high-speed projector that is expensive and is not adequate for actual applications. On the other hand, we aim to develop dynamic projection mapping system focusing actual application such as Noh performance, dance performance, etc. by using off-the-shelf equipment.
- (2) In actual stage performances, it is natural that multiple performers achieve the performances. By preparing multiple depth-sensors and by merging the obtained results, we try to detect and trace the position and shapes of multiple performers. Also by using multiple projectors, we try to achieve 3D projection mapping on multiple performers avoiding the problem of occlusion.

4 Structure of the System

4.1 Basic Structure of the System

The basic concept of our dynamic projection mapping system on non-rigid moving objects is illustrated by Fig. 5. As the dynamic projection mapping on non-rigid moving objects is the target of the topic, we neglect the projection mapping on background wall and focus on the projection mapping on multiple moving performers.

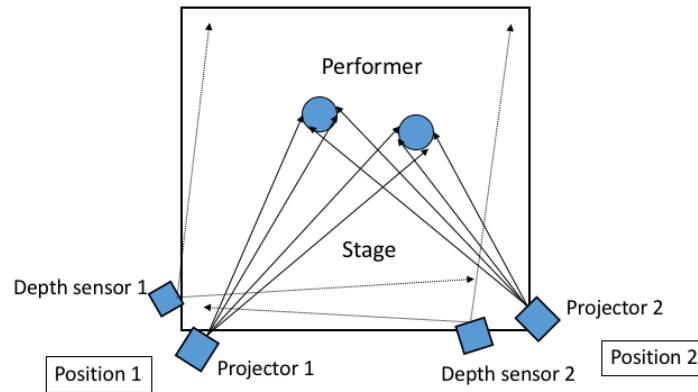


Fig. 5. Conceptual illustration of projection mapping system on non-rigid moving objects.

It is hypothesized that the width of the stage would be 6m x 6m (the size of Nor stage is 6m x 6m). Two depth sensors and two projectors are set at the two edge positions of the front side of the stage (position 1 and position 2). In our first trial we tried to set a projector and a sensor at the same place as much as possible so that the calibration between the sensor and the projector is not necessary. However, in our setting we have decided that there should be freedom for the position of depth sensors and projectors as we have introduced calibration process. It is considered that this is the basic unit system and when the stage is wider than 6m, multiple basic unit system will be used to cover the whole stage.

The sensors 1 and 2 detect and trace the positions and shapes of multiple performers and by merging the detection results the 3D models of multiple performers are constructed. Then the images to be projected on the performers are rendered on the 3D models and using multiple projectors the projection mapping is to be carried out. By using a single sensor, the obtained model is not complete 3D model. Also, as there are multiple performers, the problem of occlusion frequently occurs. The usage of multiple sensors will solve these problems to satisfactory extent. Also, in the projection stage, although the projection on a performer by the projector 1 is achieved only on the front side of the performer, projector 2 can do the supplementary role by projecting on the performer on different side of it, providing satisfactory projection result to the viewers. Figure 6 illustrates the structure of the software of the system.

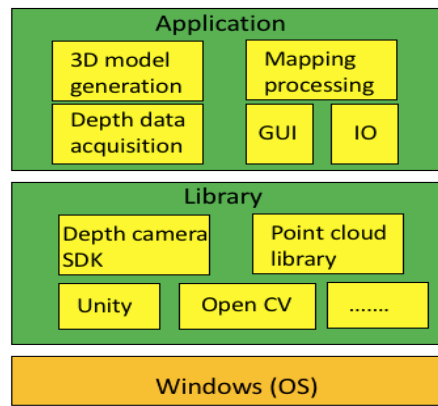


Fig. 6. Construction of the software of the projection mapping system

4.2 Tracking of Moving Objects Using Depth Sensors

The process flow of the detection and tracking of moving objects using depth sensors is illustrated in Fig 7. Below are several explanation of the each process.

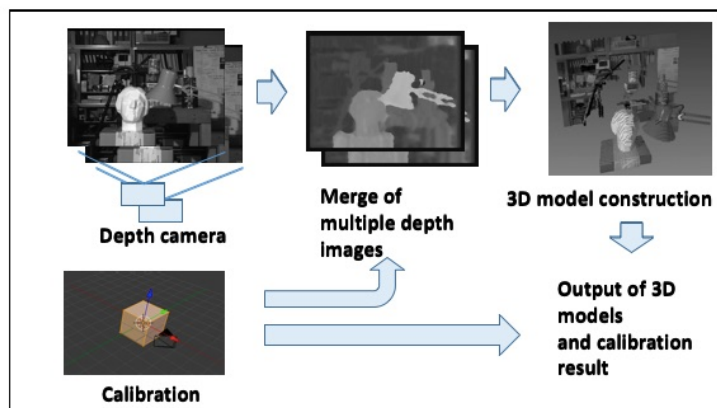


Fig. 7. Fig. 7 Process flow of the detection and tracking of moving objects

Calibration:

As the positions of two depth sensors and two projectors are different, calibration among these devices is necessary to be carried out. As a first step, calibration among the projector 1 and the depth sensors 1, 2 is carried out. The calibration is carried out by projecting several grid patterns on some object by the projector 1 and by detecting the patterns by the cameras mounted on each of the depth sensors 1, 2. Also the calibration among the projector 2 and the depth sensors 1, 2 is carried out by the projector 2 projecting grid patterns and by the cameras of the depth sensors 1, 2 detecting the patterns. These processes realize the calibration among each of the four devices, the depth sensors 1, 2 and the projectors 1, 2.

Depth image acquisition by multiple depth sensors:

Then two depth images, depth image 1 and depth image 2, of the moving objects in front of the depth sensors are obtained by the depth sensor1 and depth sensor 2.

For the depth sensor we used Kinect. However, instead of using the software development kit provided by Microsoft, we developed the necessary software by ourselves. By doing so we found that the limitations of Kinect, distance limitation and angle range limitation, is largely reduced. Now the device can detect object within 8m distance instead of original 3m capability. Also for angle range, objects within 90 degree range can be detected instead of original 60 degree range.

Merge of multiple depth images:

Then these two depth images are merged. For the depth image 2, based on the result of the calibration, it is converted the depth image 2', which is the depth image obtained at the position of the depth sensor 1, and then the two images are merged. As the depth image obtained by the depth sensor 2 covers the area that cannot be seen by the depth sensor 1, this merge will make the obtained depth image more precise.

3D image construction

The depth data, which is a set of points data with its position in the world coordinate system and distance from the depth sensors, is considered to construct the 3D image of the surroundings. The obtained 3D image is sent to the next stage for the 3D projection mapping.

4.3 Projection Mapping on Moving Objects.

The process flow of the projection mapping on moving objects and tracking is illustrated in Fig 8. Below are several explanation of the each process.

Object detection and conversion

From the set of depth data, depth data set corresponding to the moving objects to be tracked are detected. For the detection the distance between each depth data and the depth sensor is used for the key. The obtained depth data set is considered as 3D model data of the objects to be tracked and to be projected. Then this 3D model is converted into two models; 3D model 1 is the model of the object seen from the position of the projector 1 and 3D model 2 is the model from the position of the projector 2.

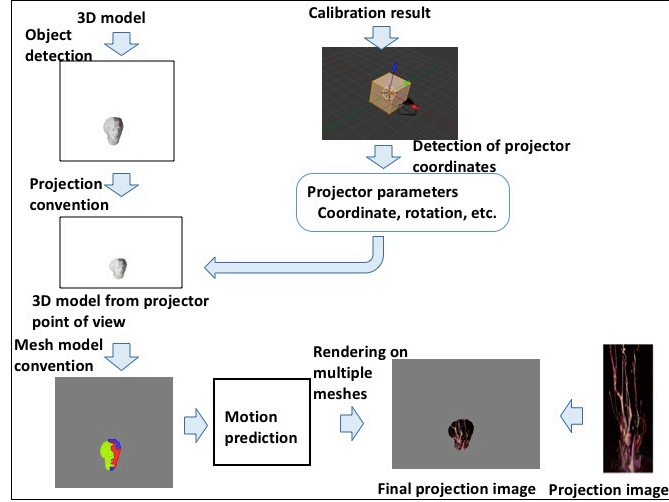


Fig. 8. The flow of the 3D projection mapping on moving objects.

Mesh model conversion

Still at this stage each model is a set of depth point data. From these point data relevant mesh data that construct the surface the objects are determined. As the number of the meshes greatly influences the calculation time of the rendering described below, the number of the meshes is to be determined carefully taking the PC's calculation capability into consideration. As the model 1 and model 2 provide supplementary 3D model for the objects, the meshes for the projection 1 should focus on the planes situated in front of the projector 1 and for projector 2 vice versa.

Motion prediction

While our system should run at real-time frame rates, the object tracking suffers from delays introduced by the round trip time through our computation pipeline for tracking. This is noticeable as a lag between the actual position and shape of a performer and the projected image. Since this effect impairs the immersion for the user, motion prediction is to be implemented in our system. We use a statistically based predictor to fit a spline into the past N frames that is extrapolated to predict the tracking trajectory. An additional weighted low-pass filter over new frames improves the robustness against outliers.

Rendering of projection images and output

Finally, the image to be projected on the object are rendered on each of the detected meshes both for the projector 1 and the projector 2. Then the final projection is carried out using the projectors 1 and 2.

5. Conclusion

In this paper we have proposed the projection mapping system that can project 3D images onto multiple non-rigid moving objects. Although projection mapping has become a new image/video projection method for entertainment and other applications,

most of the projection mapping so far proposed have the function of projecting 3D images on objects such as buildings that do not move and do not change their shapes. On the other hand, if projection mapping on non-rigid moving objects such as Noh performers or dancers is realized, the application area of projection mapping will become far wider.

To confirm this assumption, as a first step we have carried out a simple experiment in which the integration of projection mapping and Noh performance, one of the most well known Japanese traditional stage performance, was carried out. In the experiment two types of video images were projected while Noh performance was carrying out. Video images that suggest the Noh story development was projected on the back wall of the stage. In addition to this, 2D video images that suggest inner emotion of a Noh main character was projected on the Noh performer. Based on the interview-based evaluation of the system, viewers of the performance, especially most of the international students who viewed the performance, replied that, although the projected images were only 2D the produced effect was very immersive and also the Noh performance became easier to understand. Based on this experiment, we came to the belief that the integration of stage performance such as Noh and projection mapping has the possibility of creating new types of performing arts.

Based on this, we have proposed the concept of a projection mapping system that can trace the positions and shapes of moving object and can achieve 3D projection mapping on the moving objects, and developed the first prototype of the system. From hardware aspect, the system consists of multiple depth sensors and also multiple projectors and PCs. From software aspect the system consists of two parts; the objects detection part and the 3D projection part. In the object detection part, multiple depth sensors detect and trace the positions and shapes of multiple moving non-rigid objects such as dancers or performers, compose 3D models for each of the objects and give them to the projection part. Because of the usage of multiple depth sensors, even if multiple objects are moving around such as in dance performances, precise detection of positions and postures of the multiple moving objects became possible. Then in the projection mapping part, using the obtained 3D model of the objects, the object is represented as a mesh model, in which each mesh is a plane, and the images to be projected on these moving objects are rendered on these meshes and then projected on the moving objects.

The system is now being brushed up and will be used for a Noh performance as a first application of the system. The demonstration video will be shown at the time of the paper presentation.

References

- [1] Murayama, S., Torii, I., Ishii, N., "Development of Projection Mapping with Utility of Digital Signage," IIAI 3rd International Conference on Advanced Applied Informatics, pp.895-900 (2014).
- [2] Rowe, A., "Designing for engagement in mixed reality experiences that combine projection mapping and camera-based interaction," Digital Creativity, Vol.25, No.2, pp.155-168 (2013).
- [3] Bimber, O., Raskar R., "Spatial Augmented Reality; Merging Real and Virtual Worlds," Boca Raton, FL, USA: CRC Press (2005).
- [4] Mine, M., Rose, D., Yang, B., van Baar, J., Grundhofer, A., "Projection-based augmented reality in Disney theme parks," Computer, Vol.45, No.7, pp.32-40 (2012).

- [5] Maniello, D., "Augmented Reality in Public Spaces Basic Techniques for Video Mapping Volume 1," Paperback (2015).
- [6] Tosa, N., Nakatsu, R., Yunian, P., "Creation of Media Art Utilizing Fluid Dynamics," 2017 International Conference on Culture and Computing, pp.130-135 (2017).
- [7] Yunian, P., Liang, Z., Nakatsu, R., Tosa, N., "A Study of Variable Control of Sound Vibration Form (SVF) for Media Art Creation," 2017 International Conference on Culture and Computing, pp.137-142 (2017).
- [8] Toshinobu Yasumura, "Rimpa: Decorative Japanese Painting," PIE Books (2011).
- [9] Tosa, N., Nakatsu, R. Pan Yunian, "Projection Mapping Celebrating RIMPA 400th Anniversary," 2015 International Conference on Culture and Computing (2015).
- [10] Sadler, A. L., "Japanese Plays: Classic Noh, Kyogen and Kabuki Works," Tuttle Publishing (2011).
- [11] Sandra R., Raskar, G., Welch, K.-L. Low, and Bandyopadhyay, D., "Shader lamps: Animating real objects with image-based illumination," in *Rendering Techniques (Eurographics Series)*, S. Gortler and K. Myszkowski, Eds. Berlin, Germany: Springer, pp.89–102 (2001).
- [12] Bimber O., Iwai, D., Wetzstein, G., Grundhofer, A., "The visual computing of projector-camera systems," *Comput. Graph. Forum*, Vol.27, No.8, pp.2219–2245 (2008).
- [13] Fua, P., Leclerc, Y. G., "Object-centered surface reconstruction: Combining multi-image stereo and shading," *Int. J. Computer Vision*, Vol.16, No.1, pp.35–56 (1995).
- [14] Gay-Bellile, V., Bartoli, A., Sayd, P., "Deformable surface augmentation in spite of self-occlusions," in *Proc. 6th IEEE ACM Int. Symp. Mixed Augmented Reality*, pp.235–238 (2007).
- [15] Uchiyama, H., Marchand, E., "Object detection and pose tracking for augmented reality: Recent approaches," in *Proc. 18th Korea-Japan Joint Workshop Frontiers Comput. Vision*, pp.1–8 (2012).
- [16] Park, H., Park, J.-I., "Invisible marker based augmented reality system," in *Proc. SPIE Visual Commun. Image Process.*, Art. No.59601 (2006).
- [17] Guskov, I., "Efficient tracking of regular patterns on non-rigid geometry," In *Proc. 16th Int. Conf. Pattern Recognition*, Vol.2, pp.1057-1060 (2002).
- [18] Narita, G., Watanabe, Y., Ishikawa, M., "Dynamic Projection Mapping onto Deforming Non-Rigid Surface Using Deformable Dot Cluster Marker," *IEEE Trans. On Visualization and Computer Graphics*, Vol.23, No.3, pp.1235-1248 (2017)
- [19] Jones, B. R., Benko, H., Ofek, E., Wilson, A. D., "Illumiroom: Peripheral projected illusions for interactive experiences," In *ACM SIGGRAPH 2013 Emerging Technologies*, ACM, New York, NY, USA, SIGGRAPH '13, 7:1–7:1 (2013).
- [20] Siegl, C., et al., "Real-time pixel luminance optimization for dynamic multi-projection mapping," *ACM Trans. Graph.*, Vol.34, No.6, pp.237:1–237:11 (2015).