



Virtual Participation in Ukiyo-e Appreciation using Body Motion

Shota Kusajima
Future University Hakodate
s-kusajima@sumilab.org

Takuya Takahashi
Future University Hakodate
t-takahashi@sumilab.org

Yasuyuki Sumi
Future University Hakodate
sumi@acm.org

ABSTRACT

This paper presents a novel method of art appreciation by participating in artwork using virtual reality technology. We chose a famous ukiyo-e work titled “many people are assembled to create one good person” drawn by Kuniyoshi Utagawa in the Edo period of Japan. The painting provides an illusion that it seems a one man but shows many (over ten) persons assembled to create the one person. We developed a 3D CG game where its player can join to create the big person by virtually participating into the artwork as the part of the big person. The system enabled its players to enjoy the game with their own body movements, which extends their experience of art appreciation. This paper describes an attempt to enrich our game players’ experience under very limited temporal and spatial constraints given at a contest exhibition.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; **Empirical studies in interaction design**; • **Applied computing** → **Performing arts**;

KEYWORDS

Virtual reality, Art appreciation, Bodily kinesthetic game, Content design

ACM Reference Format:

Shota Kusajima, Takuya Takahashi, and Yasuyuki Sumi. 2018. Virtual Participation in Ukiyo-e Appreciation using Body Motion. In *AH2018: The 9th Augmented Human International Conference, February 7–9, 2018, Seoul, Republic of Korea*. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3174910.3174946>

1 INTRODUCTION

“Yose-e” is a genre of Japanese ukiyo-e paintings that provides an illusion that it seems a one big person or object but shows many people or objects assembled to create the big person/object. Giuseppe Arcimboldo is world-widely famous for such intricate paintings, which combined inanimate or found objects into a portrait that would resemble the portrait subject. Similarly, Japanese ukiyo-e painters in Edo period drew “yose-e” representing various subjects by combining smaller components.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

AH2018, February 7–9, 2018, Seoul, Republic of Korea

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

ACM ISBN 978-1-4503-5415-8/18/02...\$15.00

<https://doi.org/10.1145/3174910.3174946>

There has been attempts to create such a collage image from many images helped by computer graphics techniques [1][2]. However, these paintings are imaginary, so it is physically difficult for an actual human to try to reproduce like group gymnastics as a member of yose-e.

We created an experiential content in which a player can experience poses taken by yose-e members in the virtual reality space, and aimed for a new viewing experience of paintings. In the content, we chose a famous ukiyo-e work titled “many people are assembled to create one good person” drawn by Kuniyoshi Utagawa (Figure 1). Through this VR content, it was expected that the players enjoy the difficulty and absurdity of the poses taken by members, and a new finding is facilitated by spatially appreciating paintings drawn on a flat surface.

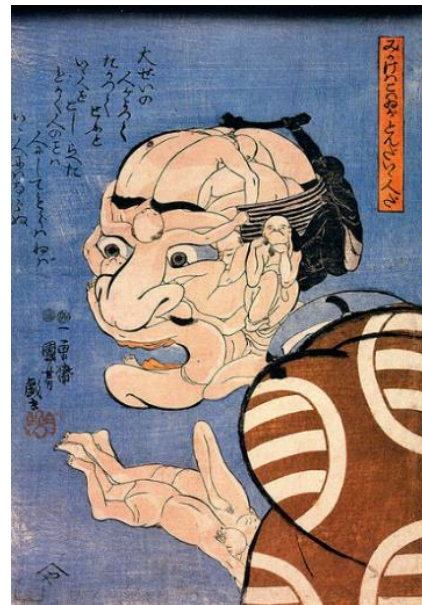


Figure 1: “Many people are assembled to create one good person” by Kuniyoshi Utagawa

Traditional experiential contents including VR games can be enjoyed only by players. However, we had hard constraints with very limited time and space at the contest venue where we exhibited our system, so it was difficult to realize multiple players simultaneously enjoy the game. We wanted to avoid the situation where only the game players enjoyed and others waiting in a long queue became bored. Hence, we aimed to develop a game content that audiences watching the game play of others can also enjoy.

In this paper, we describe spatial and temporal design of interactive contents through the composition and production of experiential content, and the results of the experience exhibition.

2 RELATED WORK

2.1 3D CG animation and collage

There has been attempts to synthesize 3D CG objects to create an artistic collage [3, 4]. However, these attempts combine just existing images, not to create new components during combining. In order to make it experience-based content, it is necessary to handle things generated by players on the fly. Some researches measuring movement of people have been done in 90s [5], and recently it is also possible to create models that match the movement of a person's body in real time [6–8]. In this research, by emphasizing real-time nature using these techniques, we create a content that can encourage active participation of players.

2.2 Reproduction and immersion of works

There are many masterpieces that are difficult to reproduce. In the other hand, it is easy to reproduce such works digitally, and so such attempts have been made [9, 10]. By doing this, we can browse easily and safely, digital reproduction is widely adopted. In addition to paintings and art objects, there are things that reproduce the entire city in virtual space, and it makes us feel like actually visited [11]. With the development of VR technology, it became possible to actually experience themselves as if they are on the spot and touches works and landscapes. In addition to being able to appreciate painting work using VR technology, we aim to participate in the work and aim to connect to a new viewing experience through creation of original work.

2.3 Interactive contents

Our system is a game where its player participates in interactive content creation by controlling CG character in 3D virtual space by their own body motion. Some similar works have been proposed by capturing the player's body motion by RGB-D camera (e.g., Kinect) in the context of creating cartoon-like pictures [12, 13]. We also employed similar game design that a player is asked to pose to complete missing parts of the artworks, but extended the idea from 2D image to 3D space interaction.

3 CONTENT DESIGN

3.1 Content overview and constraints

In this content, a player takes poses of yose-e members, and makes player's own yose-e by assembling these poses like puzzle pieces. Figure 2 is an image diagram in which player's own pose becomes part of a painting.

In the content, the player is instructed to take poses and reproduce the poses using the whole body. When all the poses instructed have been reproduced, a 3D model of yose-e made on the basis of player's poses and a 3D model based on the original painting are displayed in the VR space. In addition, some 3D models created by other players in the previous plays are also placed around the player. The player mounting head mount display can watch 3D models from various places and angles. An audience can enjoy the



Figure 2: Conceptual image of virtual participation in ukiyo-e

absurdity trying to imitate the poses of yose-e members in whole body. They can enjoy with the player to give an instruction of posture.

This content was created as a VR work for exhibiting in International collegiate Virtual Reality Contest (IVRC) held in September and October 2016. In this contest, the exhibition place was restricted 1.8 meters square, so we created the content to satisfy this restriction.

3.2 Used equipment

In this content, we used Kinect v2 to measure the movement of the player taking poses of yose-e member. As a method of measuring player's motion, a motion capture system or the like can be considered, but this time we decided to use Kinect v2 because it takes time and effort to attach tracking markers and we cannot secure a large space. Also, the player mounted an Oculus Rift to watch the game screen and created 3D models from various angles and places. A human model is displayed on the play screen, and moves in the same way in conjunction with the player by posing in front of Kinect v2. The player takes poses and creates a piece of the puzzle while checking that model. Content creation using Oculus Rift and Kinect has been done so far [14], by using Unity which is a multi-platform game engine, it is easy to reflect player's pose in virtual 3D space.

Owing to the restriction that the place where content can be exhibited was 1.8 meters square, we placed Kinect at the corner and put the player's position on the diagonal. Because it was thought that some players take crouching or sitting poses, Kinect was placed to look down from diagonally above in order to measure player's motion with one Kinect. Kinect was installed at a height of 1.8 meter from the ground where the standing player.

Also, a display was placed outside the playing space. On this screen, the same image as the VR space which the player was watching was displayed. The audience watching the stage of play from the outside can enjoy the content with the same line of sight as the player.

Figure 3 and 4 are schematic diagrams of the exhibition space of content.

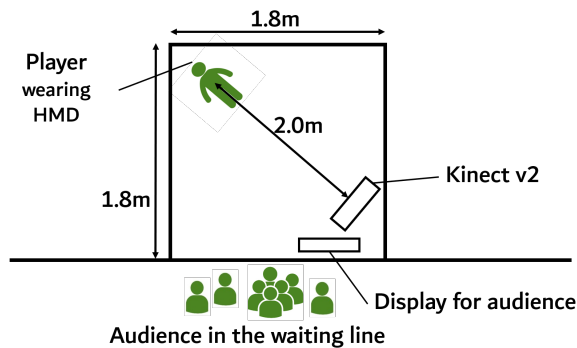


Figure 3: Schematic diagram of the exhibition space (overview)

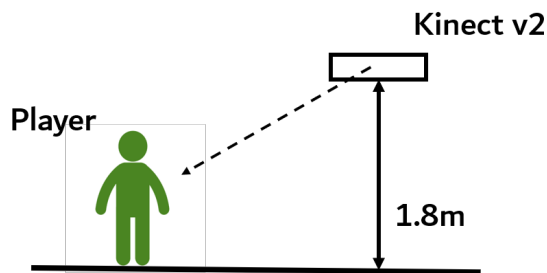


Figure 4: Schematic diagram of the exhibition space (side-view)

3.3 Composition of the content

This content is roughly divided into three parts: explanation, experience, and appreciation. Figure 5 is a flow diagram of this content. In the following, we describe the outline of each part and design points.

3.3.1 Explanation part. In the first part, the player receives explanations about creating the pieces of yose-e by player's motions and part of the pose the player takes in about 20 seconds. The stage of the game is conscious of the streets of Edo in Japan, which is the era when the yose-e was drawn. Here, two people in Edo instruct the poses to the player as navigators (Figure 6).

At the start, human models making up yose-e are displayed on the screen. The human model standing at the front has the same behavior according to the movement of the player. In this way, without being explained deeply, the player can understand how to move the body and how to operate.

3.3.2 Experience part. In the following part, the player creates parts of yose-e by moving the body. The original painting is drawn as a combination of more than a dozen people. But in this content, we limited the poses the player takes to the ear, nose, and lips with emphasis on ease of play to one time and ease of posing. The player creates three poses within the time limit (Figure 7). We decided to take 10 seconds to spend taking each pose in order to be conscious of posing intuitively. Also, we made a flow to take from easy pose to difficult one. The player makes an easy ear pose at first to check own behavior. Next the player makes a nose pose that is in the front and has to move the body exaggeratedly when find out the operation to some extent. Finally, the player creates lips that are difficult to take a pose.

At the beginning, a 3D model of the whole face with all parts is displayed. After that, when the player takes poses of ear, nose and lips, each part is hidden. Because players take poses without looking at the correctly shaped parts, it is expected that diversity will be born, taking poses resembling a face, or taking ingenious poses. When taking poses, since the human model moves to position of the part taking a pose, the player can take poses without moving considerably from the home position. With this, it is possible to play even in a narrow space.

3.3.3 Appreciation part. In the last part, the 3D model added poses taken by the player and the 3D model created based on the work of Kuniyoshi Utagawa are displayed side by side (Figure 8). These models are located in the VR space, so the player with HMD can appreciate them from various directions (Figure 9). This allows the player to see details. Not only that, the player compares them and finds differences. The player can see the posture of human that are difficult to imagine just by looking at the original paintings.

In addition, other models created by other players in the previous plays are also displayed around the player (Figure 10). Thereby, the player can enjoy the difference between other works and the player's, and acquire a new perspective on the paintings.

3.3.4 After playing. After play, the player can look back on what kind of pose while playing. We take the image of the player while taking either pose of ear, nose or lips with Kinect v2, and superimpose the finished 3D face models on the face part of the image (Figure 11). Finally, by reducing the color of the image, an image like ukiyo-e is created. By looking back the image, the player can see how the poses become parts of ukiyo-e, and it is expected to lead to a new viewing experience.

4 EXPERIENCE EXHIBITION IN THE CONTEST

We exhibited the VR content in a preliminary contest, 24th International collegiate Virtual Reality Contest on the 15th and 16th September 2016, and in a final contest on the 29th and 30th October 2016. The two contests were open to the public, and approximately 250 visitors played the content.

4.1 State of experience exhibition

In the preliminary contest, approximately 100 visitors, and in the final contest, approximately 150 visitors experienced our content. Not only the players but also the audience were enjoying it to be

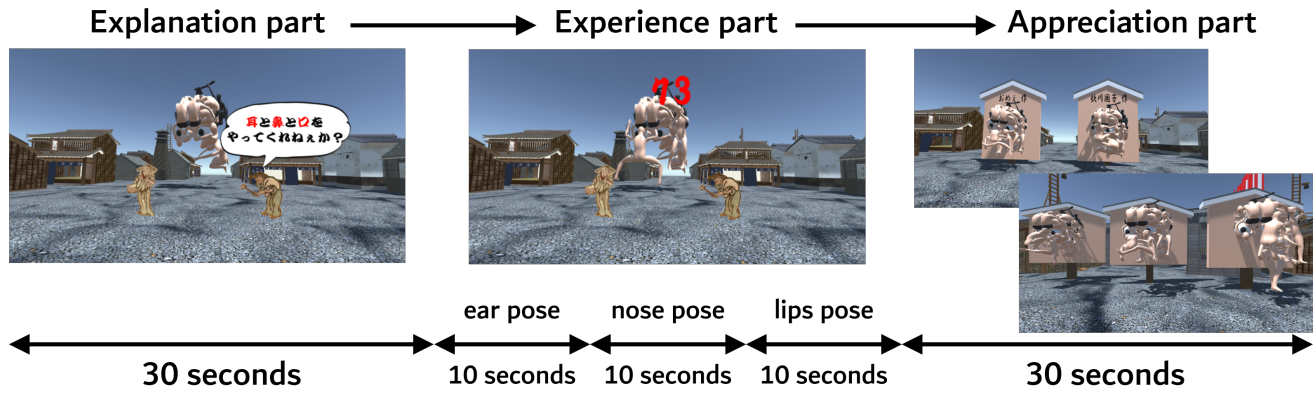


Figure 5: Flow of gaming experience



Figure 6: Opening scene of the game: Invitation to the game

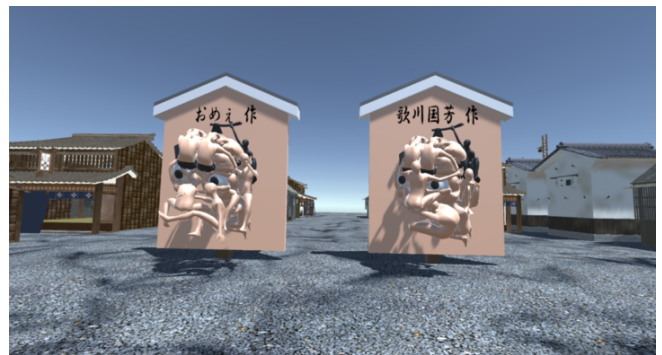


Figure 8: Work result by the player and model based on the original ukiyo-e



Figure 7: Physical experience of posing (of nose)



Figure 9: View from below

interesting in the absurdity of the players posing for the members of the yose-e which are difficult to reproduce. Many players were laughing at the absurdity of their own poses which they were taking while playing to look at an image in which a finished 3D face model was superimposed on the pose that they were taking. Figure 12, 13 and 14 show the state of playing by a visitor.

The yose-e used as puzzles were rather famous, but we expected that many people did not know the paintings or the poses of people in paintings. We expected various types of poses were seen for each of the ear, nose and lips because we did not tell the player which pose to take in advance. However, regarding the ear and nose poses,

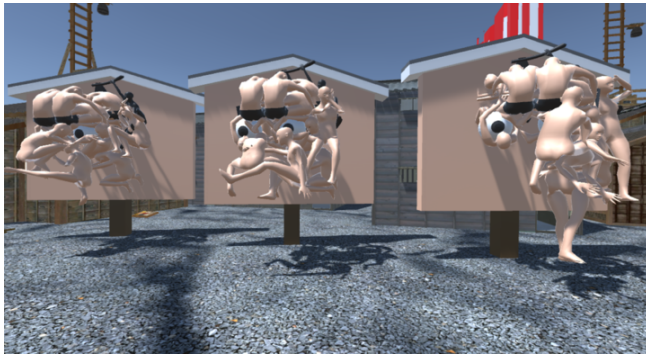


Figure 10: Works by other players



Figure 11: Ukiyo-e like image with superimposing the work on player's pose

they were classified into some kinds of types. Regarding the ear poses, it was able to be divided into two patterns.

- expressing the ears using only the arm
- expressing with the whole body bending the body to the letter "C"

Regarding the nose poses, it was also able to be divided into two patterns.

- spreading arms with crouching
- fold the waist while turning sideways

On the other hand, the lips poses were the most difficult, and the moving model was facing sideways, so it was difficult for players to reproduce. Regarding these poses, various patterns were seen,



Figure 12: Appearance of exhibition site



Figure 13: Player and display showing his view

for example, expressing with both arms, or sitting on the ground and representing the lips with their legs.

In the appreciation part, it was expected that a new reaction could be obtained by placing the 3D models created by other players around. Actually, there were the following reactions.

- evaluation on completeness of other people's works
- evaluation of entertainment and absurdity for works whose reproduction level was not so high
- reproducing poses taken by other players
- comparison of differences from the player's work
- laughing at the absurdity of the player's work

From these reactions, it turned out that the player tended not only to make a work of high degree of completion, but also to have fun of making works featuring only player's own.



Figure 14: Player viewing his work from below

4.2 Design considerations

In this design, it was necessary to create a place of playing in a narrow space of 1.8 meters square due to restrictions on the contest. For that reason, we were not able to install a large-scale system such as motion capture, and to let multiple players play at the same time. However, by players taking various humorous poses, we were able to make it possible to enjoy the audience watching playing. It was very important in exhibition type contents to take a configuration that only people playing can enjoy the contents, but also audience who are viewing the playing can enjoy. If there are no spatial constraints, the accuracy and diversity of poses taken by players will increase by using motion capture system. Also, as more than one player can participate at the same time, it is expected to become more developed content.

On the temporal design aspect, we tried to shorten the time by making it possible to understand how to move the body and the purpose of the game through playing without detailed advance explanation of operation in order to have many participants experience in time. Especially, we set the time to get each pose to 10 seconds each as short time. The players took various poses while devising ingenuity in a short time, and it seemed that the audience was also enjoying it by watching the plays. Also, by raising the degree of difficulty gradually from simple pose to difficult pose, the players got accustomed to the operation method even with few explanations and a short time, and it was able to proceed well.

5 CONCLUSION

In this paper, in order to realize a new viewing experience of painting works, we made an exhibition that reproduces a 2D painting drawn by Kuniyoshi Utagawa with the poses taken by players in a 3D VR space. Some players made works close to the original work taking poses that resembles the face parts, and others made unique works by daring to take different poses than the real one. At the same time, we were able to make it possible for not only the players but also the audience watching plays to enjoy the difficulty of

reproducing the poses taken by the members of actual painting, and the absurdity of being able to reproduce strange poses.

By having many people experience the content, we were able to consider the content design from the reactions of experiences and viewers to the content. In creating contents, there are various restrictions including time and space. It is necessary to find a suitable design from each constraint, and we think we were able to find something to be kind of indicator in the making like this time.

ACKNOWLEDGMENTS

We thank Ayuri Tomohiro, Junki Kikuchi and Seiya Hamamoto for helping the implementation and exhibition of the system described in this paper.

REFERENCES

- [1] Hua Huang, Lei Zhang, and Hong-Chao Zhang. Arcimboldo-like collage using internet images. In *Proceedings of the 2011 SIGGRAPH Asia Conference*, pages 155:1–155:8, New York, NY, USA, 2011. ACM.
- [2] Cheng-Han Yang, Ying-Miao Kuo, and Hung-Kuo Chu. Synthesizing emerging images from photographs. In *Proceedings of the 2016 ACM on Multimedia Conference*, pages 660–664, New York, NY, USA, 2016. ACM.
- [3] Ran Gal, Olga Sorkine, Tiberiu Popa, Alla Sheffer, and Daniel Cohen-Or. 3d collage: Expressive non-realistic modeling. In *Proceedings of the 5th International Symposium on Non-photorealistic Animation and Rendering (NPAR '07)*, pages 7–14, New York, NY, USA, 2007. ACM.
- [4] Asier Marzo and Oscar Ardaiz. Collart: A tool for creating 3d photo collages using mobile augmented reality. In *Proceedings of the 21st ACM International Conference on Multimedia*, pages 585–588, New York, NY, USA, 2013. ACM.
- [5] Christopher Wren, Ali Azarbayejani, Trevor Darrell, and Alex Pentland. Pfnder: Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:780–785, 1997.
- [6] Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, and Andrew Fitzgibbon. Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST '11)*, pages 559–568, 2011.
- [7] T. H. D. Nguyen, T. C. T. Qui, K. Xu, A. D. Cheok, S. L. Teo, Z. Y. Zhou, A. Mallawaarachchi, S. P. Lee, W. Liu, H. S. Teo, L. N. Thang, Y. Li, and H. Kato. Real-time 3D human capture system for mixed-reality art and entertainment. *IEEE Transactions on Visualization and Computer Graphics*, 11(6):706–721, Nov 2005.
- [8] Jamie Shotton, Toby Sharp, Alex Kipman, Andrew Fitzgibbon, Mark Finocchio, Andrew Blake, Mat Cook, and Richard Moore. Real-time human pose recognition in parts from single depth images. *Communications of the ACM*, 56(1):116–124, January 2013.
- [9] Shoji Yamada. Who moved my masterpiece? digital reproduction, replacement, and the vanishing cultural heritage of kyoto. *International Journal of Cultural Property*, 24(3):295–320, 2017.
- [10] Michael Fischer. A proposal for a virtual reality museum for virtual reality art. In *Proceedings of Museums and the Web 2017 (MW 2017)*, 2017.
- [11] Tomoki Nakaya, Keiji Yano, Yuzuru Isoda, Tatsunori Kawasumi, Yutaka Takase, Takashi Kirimura, Akihiro Tsukamoto, Ayako Matsumoto, Toshikazu Seto, and Takafusa Iizuka. *Virtual Kyoto Project: Digital Diorama of the Past, Present, and Future of the Historical City of Kyoto*, pages 173–187. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
- [12] Daiki Umeda, Tomoaki Moriya, and Tokiichiro Takahashi. Real-time manga-like depiction based on interpretation of bodily movements by using kinect. In *SIGGRAPH Asia 2012 Technical Briefs*, pages 28:1–28:4, New York, NY, USA, 2012. ACM.
- [13] Yuto Nara, Genki Kunitomi, Yukua Koide, Wataru Fujimura, and Akihiko Shirai. Manga generator: Immersive posing role playing game in manga world. In *Proceedings of the Virtual Reality International Conference: Laval Virtual (VRIC '13)*, pages 27:1–27:5, New York, NY, USA, 2013. ACM.
- [14] Dongjik Lee, Giyeol Baek, Yangwon Lim, and Hankyu Lim. Virtual reality contents using the oculus rift and kinect. *Proceedings of the 2nd International Conference on Mathematics and Computers in Sciences and Industry (MCSI 2015)*, pages 102–104, 2015.