ISSN 1738-8074

International Journal of Asia Digital Art and Design

Vol.18 No.02 2014
Asia Digital Art and Design Association

Contents

Categories for paper

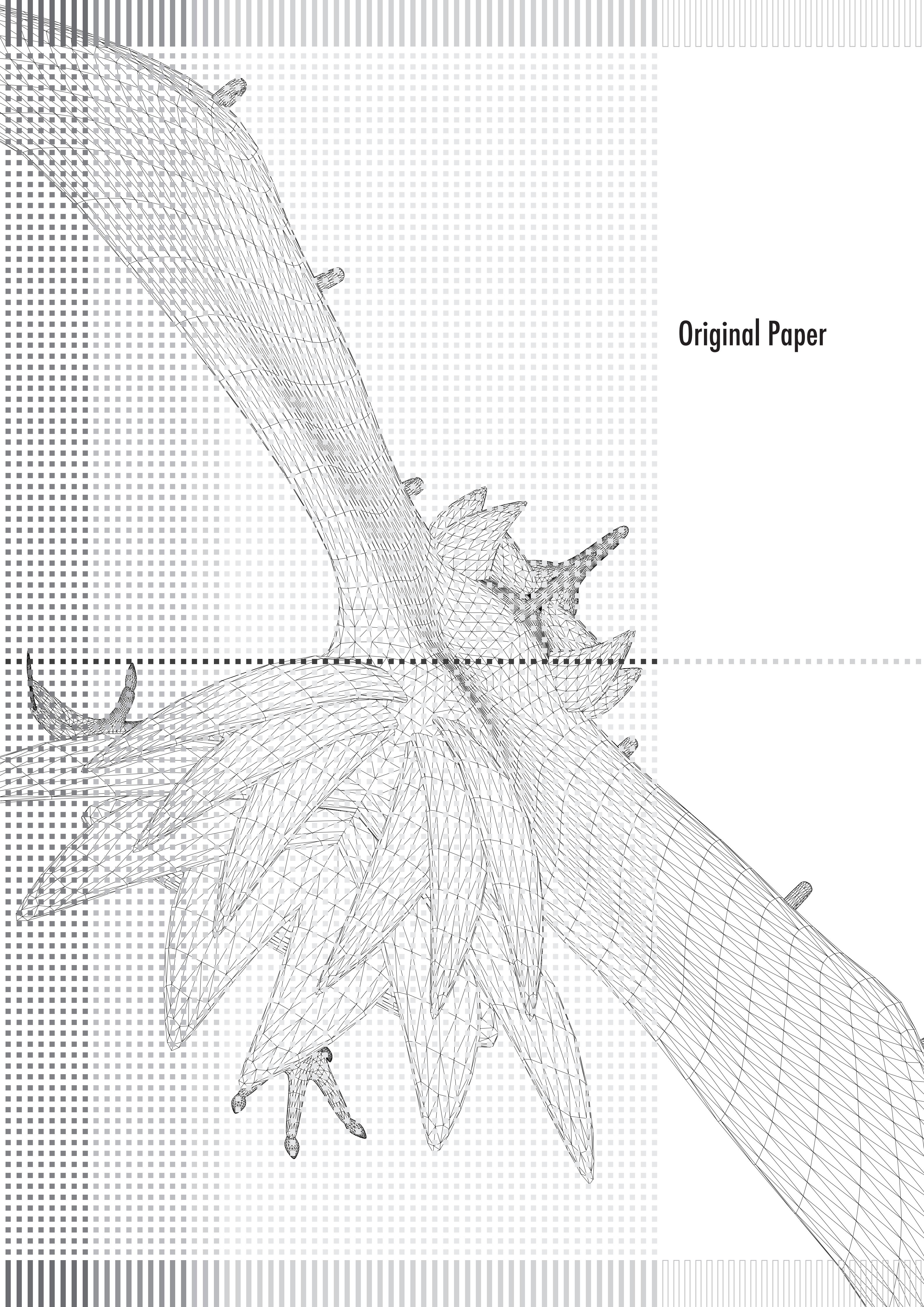
・Original Article

A paper in this category has to be a logical and empirical report of the study, the review and the proposal by the author on the issue of digital art and design based on media technology. It also has to include the novelty and academic values which can be shared with ADADA members or the people who study digital art and design.
Number of pages: 6 -10

・Art Paper

A paper in this category has to consist of the author’s practice, result and expository writing on the field of digital art and design. It also has to have the originality in its concepts, methods, expression techniques and making process, and the result should have some values which can be shared with ADADA members or the people who study digital art and design.
Number of pages: 6 -10

Original Paper



Received April 6, 2014; Accepted June 15, 2014

International Journal of Asia Digital Art & Design

Effect of Scene Emphasis by Pseudo Displacement of Viewing Distance in Stereoscopic Image

Tatsuro, Ishii.
Kyushu University
tatsuro@design.kyushu-u.ac.jp

Dae-Woong, Kim.
Kyushu University
dwkim@design.kyushu-u.ac.jp

Abstract

In a three-dimensional vision, as for the picture exhibited by a monitor or a screen, size and the aspect ratio of the images are usually fixed. However, we think that scenes can be visually emphasized more by changing the presentation position of an image, i.e., the position of a screen, or a viewing angle. In this research, we examined the displaying method that Pseudo-screen is placed in the Stereoscopic 3D (S3D) space. Pseudo-screen induce changing the viewing distance artificially. And pseudo-screen itself displays S3D images. We verified whether the scene would be emphasized, in other words, the sense of "watching it nearby". The prior study or displaying method about previously described are uncommon.

As a first step, we evaluated visual pattern of the pseudo-screen and surrounding by the impression evaluation experiment for students. At the second step, we displayed the visual stimulus, in which the position of pseudo-screen changes with the field of view of shots. These shots constituted a scene as montage. In the examination, we verified whether the stimulus caused change of viewing distance, emphasis of the scene, and is recognized as montage. Regarding the pattern evaluation, the images that wireframe or plane is extended from edge of the pseudo-screen are suitable for the display method. And, in the scene display examination, the shots constructed simple wireframe or plane are evaluated. Most participants recognized that the position of the screen is changed, and "watching it nearby".

As a result, we considered that changing viewing distance caused by changing position of pseudo-screen is effective for emphasis of the scene.

Keywords: S3D image, pseudo-screen, scene emphasis

1. Introduction

When we watch a Stereoscopic 3D (S3D) images displayed on a television monitor or a screen, as well as 2D images, as for the picture exhibited by a monitor or a screen, size and the aspect ratio of the images are usually fixed. Regarding video contents, in a scene, each shots have different field of view related to the property of a camera or lens. When a subject is shot with the same visual size, the distance to a subject of camera with wide-angle lens is shorten than camera with normal lens. Because the image of the wide field of view by wide-angle lens can emphasize a sense of dynamism, it is an effective element emphasizing the impression of the shot more. In S3D image, when the image of wide angle of shot is compared with that of normal angle of shot, visual size of the subject with wide angle is smaller than the normal one, because depth information increases. In wide field of view, depth is emphasized more than normal, and this causes the miniature garden effect in S3D [1]. When it comes to S3D, the image with wide field of view which causes originally sense of dynamism, results lack in it adversely. However, we consider that the impression of the shot of the S3D image can be emphasized effectively by adopting wide angle of shot to enable dynamic expression of the point-blank range.

In this research, it is verified whether the sense of "watching at a S3D image at point-blank range" can be given to viewers by changing a setup of depth position of S3D image according to a viewing angle. And it is verified whether the display method is effective.

2. Background and Purpose of this research

In our research on S3D digital archives of Shinto theatrical dance (Kagura), using stereo cameras, the dedication of Kagura-dance to Shinto shrine was filmed. The "Buzen Iwato Kagura" handed down from ancient times in Buzen-shi, Fukuoka prefecture, became the Intangible Folk Cultural Properties of Fukuoka 1999, and tradition is inherited by six groups now. The group is called "Kagura-ko", and, in every autumn season, the Kagura dedication to a Shinto shrine is performed praying for a bountiful harvest flourishingly [2]. Variation of the dance "Mai" in Kagura exceed 30 kinds, and the styles of the dance are slightly different in each area, too. The Kagura dedication is performed in a stage "Kagura-den" which is built at the Shinto shrine, and it is endless, a dance is continuously performed from the early morning to the middle of the night. Local people called the parishioners "Ujiko" go to the Shinto shrine and sit in the porch of the stage and watch "Kagura dance" at close range, and many dances are performed (Fig.1). Dancers take a communication with Ujiko while they are dancing. In our research of digital archive of Kagura, S3D format video data was utilized not only as historical data, also as educational contents. And we produced S3D movie contents to contribute to local promotion and performed screening events in a museum or a public hall to spread charm of the Buzen Iwato Kagura inside and outside Fukuoka.

Fig.1 Buzen Iwato Kagura (dedicaction of Kagura)

The educational contents and S3D movie contents will help solving a problem of the tradition due to the depopulation.

We considered that the image of wide angle of shot taken at the close range from the stage is indispensable for dynamism of Kagura. Actually, viewers gave a high evaluation on the image of wide angle of shot, when we performed an interview at the screening event. However, in the present circumstances, Mixed images of the three-dimensional vision with the normal field of view and a wide field of view are displayed on the fixed size and aspect ratio of a screen (or a monitor). And some viewers claimed that they felt the visual size of subject of wide field of view is smaller than the normal. S3D image displaying, with depth information, when a wide image (taken from close range) and a normal image (taken from distance) are projected on the screen of same size and aspect ratio, the distance from a subjects to camera is changed, a sense of “watching it in close range” will not be suitable. To resolve this issue, we set up the hypothesis that changing the distance between the screen and viewers will be the effective display method to emphasize the sense of watching it at short distance for viewers. However, it is not practical to move a physical screen instantly. In this research, we examined the displaying method that a pseudo-screen is placed in the S3D space. And pseudo-screen itself displays S3D images. The pseudo-screen induce changing the viewing distance virtually. We verified whether the scene would be emphasized, in other words, the sense of “watching it nearby”

The purpose and goal of this research is to establish the displaying method for S3D image.

3. Related works, research

The study on display method to place a pseudo-screen in a S3D space is not seen. In the IMAX format S3D movie, *Encounter in the Third Dimension* (1999), there was a scene jumping into the S3D world of 3D TV monitor, and the TV monitor is placed in a room which was also displayed in S3D (Fig.2).

Here, the viewer will watch a S3D images containing two depth information at the same time in S3D space projected on physical screen, by a pseudo-3DTV monitor being placed in it,

Fig.2 *Encounter in the Third Dimension* (1999)

In this shot, the room and other elements are displayed in S3D, and the image on TV monitor is also S3D.

and a feeling of depth is emphasized more. However, it is considered that various things around the monitor cause visual fatigue, if viewer watch it for a long hours. In our method, the pseudo-screen is the image to let viewers focus on, and surrounding environment in S3D space will be minimum visual information they can grasp a sense of depth with it, thus, they won’t get visual fatigue.

Ofuji analyzed the absolute localization of viewer in S3D [3]. This study was inspection as to whether viewers got a sense of instant change of their viewing position when they were watching S3D shot replaced from other S3D shot, these shots were took from different positions. It is considered that because fixed images of the same size were projected on physical screen, the sense of changing position has been occurred to viewers. In our research, it is considered that the sense of the change of viewing position can be reduced by moving the screen (pseudo-screen) itself virtually.

4. Concept of the method

Generally, a S3D image on physical screen has one depth information, we arranged two steps of screen positioning. One screen is physical screen usually, and other one is virtual screen (**pseudo-screen**). In the space constituted from S3D image, another S3D screen (pseudo-screen) is positioned in it (Fig.3). Pseudo-screen is located various position in the space artificially. Thus, viewers will emphasize the sense of watching the (pseudo-) screen at short distance.

In the method, two Convergence Point (CP) should be set for physical screen and pseudo-screen. Controlling CPs is important for viewers to watch S3D images on pseudo-screen without incongruity. To rectify these CPs, S3D images on pseudo-screen was created first, then images of S3D space (in which pseudo-screen was placed) on physical screen was created.

Flow of the setting of the display technique is as follows.

1. Edit images of right eye and left eye for pseudo-screen, convergence point (CP) is set at the screen.
2. Create a virtual 3D space, and pseudo-screen is placed in it.
3. Using virtual stereo camera, right image is rendered with right image on pseudo-screen, left image is rendered with left image on pseudo-screen. CP is set at the physical screen.

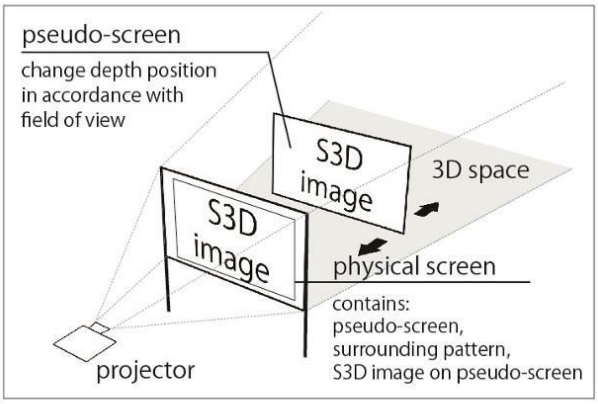


Fig.3 Concept of "Pseudo-screen in S3D space" display method

There is a possibility that viewers cannot recognize the connection of the shot as montage because the visual size changes in every shot. We thought that the connection of shots is recognizable by adding artificial depth information, the sense that the screen got closer to viewers is emphasized. In this research, we also perform the verification, that the unity of shots with different depth position will be recognized as a scene, from the view point of montage.

5. Experiment1: Display Pattern Inspection for Pseudo-screen

For "Pseudo-screen in S3D space" display method, we inspected appropriate visual pattern (pseudo-screen and surrounding element). First, we inspected suitable visual pattern of 3D space surrounding pseudo-screen. It is considered that a pattern of the line extended from the vanishing point, one-point perspective, can create the sense of depth [4]. As visual stimuli, patterns were created with wireframes and planes.

5.1 Stimuli

The S3D image displayed on pseudo-screen is the primal visual image which must be focused on. As for the visual information surrounding pseudo-screen, it is considered that too much S3D factor around pseud-screen would become hard to be watched for viewers. There are some patterns about visual information that let viewers recognize depth [4], in this study the depth expression was simply constructed by wireframe or plane, based on perspective.

Three ways of depth expression were set.

Depth expression 1: Wire frame (or plane) expands toward viewers from the edge (or the corner) of pseudo-screen as the starting point of the depth. **(LD, PD)**

Depth expression 2: Pseudo-screen is surrounded by the S3D space (in front and behind). **(EW, SP)**

Depth expression 3: Only pseudo-screen is placed in the blank space. **(NB)**

Five patterns were used as stimuli for this research. Pseudo-screen of all patterns are placed at same depth position. Convergence point was set at the physical screen. Depth position was set where visual size of pseudo-screen is 64 percent of the physical screen. Live action Kagura dance was used as a S3D image on pseudo-screen. All patterns are created using Adobe After Effects5.5 and Premier6.0, and are Full HD Mpeg2 format.

Pattern1: Extended Wireframe Type "EW"

Extension line from the structure in the picture with a wire frame (Fig 4).

Pattern2: Linear Depth Type "LD"

Wire frame starts from the corner of pseudo-screen, extend toward viewer(Fig 5).

Pattern3: Planar Depth Type "PD"

Planes are extended from the edge of pseudo-screen, toward viewer (Fig 6).

Pattern4: Space placement Type "SP"

Grid image as background is placed in S3D space, pseudo-screen is placed above it (Fig 7).

Pattern5: No Background Type "NB"

No background, the visual information is pseudo-screen only (Fig 8).

Fig.4 Pattern1: Extended Wireframe Type **"EW"**

Fig.5 Pattern2: Linear Depth Type **"LD"**

Fig.6 Pattern3: Planar Depth Type **"PD"**

Fig.7 Pattern4: Space placement Type **“SP”**

Fig.8 Pattern5: No Background Type **“NB”**

5.2 Experiment Procedure

14 students from Kyushu University participated the experiment (10 males, 4 females, all Japanese). All participants had experience to watch S3D movie. The experiment was carried out in the studio with condition of no illumination. Equipment used in the experiment are two DLP projectors (Full HD, 4000lm), circular polarizing filters, 150 inch silver screen and polarizing glasses. Mac book pro and Blacmagic Design’s Ultrastudio 3D are for playing S3D images.

Paired comparison method was used for this test. Combination all five stimuli including anteroposterior order are tested (20 pairs). For example, precedent stimulus EW (named as “A”) and following SP (named as “B”) were displayed, each stimuli was displayed 10 seconds with no sound. After paired (“A” and “B”) stimuli were displayed, a white blank image is displayed twenty seconds, participants responded to three evaluation items (“visibility” “depth” “visual size”) by five phases of evaluation standards. The evaluation sheet was notation in Japanese (Fig.9).

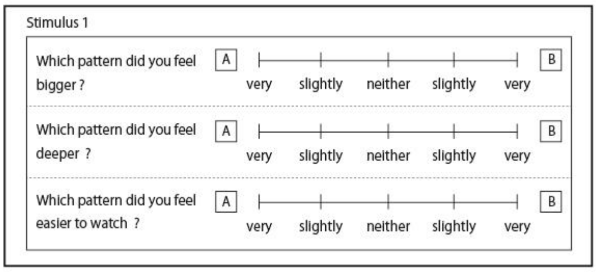


Fig.9 Evaluation sheet for experiment 1

5.3 Factor Analysis and discussion

Based on the collected data, one-way analysis of variance (factor: display pattern) and ranking by the evaluation point are performed (Table.1 and Table.2). As for the evaluation items, “visibility” and “depth”, the significant difference are confirmed (p<0.01). Regarding “visual size”, there is no significant difference. EW, LD, and PD were highly evaluated in “visibility” and “depth”. Especially as for EW, many participants got the sense of depth from building shaped image, wire frames were extended in front and backward from pseudo-screen. On the other hand, SP and NB considerably had low evaluation. As a reason, in the case of SP, screen seemed to float in the 3D space, and there were many opinions that participants had sense of incongruity. As for NB, there were many opinions that participants could not judge where pseudo-screen was located. It is considered that there is a visual difference between ED, LD, PD and SP, NB. The former are images in which wire frame or plane was extended from the edge of pseudo-screen, then participants can recognize depth easily.

Based upon the foregoing, we judged that EW, LD and PD were suitable for display method and decided to use it for the next experiment. SP and NB are excluded.

visibility ** p<0.01

Source	SS	df	F
between	6.31644	4	11.84333 **
within	5.33333	40	
total	11.64977	44	

depth ** p<0.01

Source	SS	df	F
between	5.86666	4	7.14285 **
within	8.21333	40	
total	14.08	44	

Table.1 Analysis of variance table (“visibility” and “depth”)

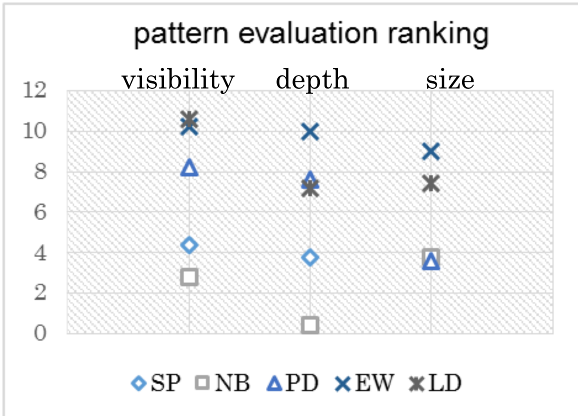


Table.2 ranking by evaluation point

6. Experiment2: Scene Effectiveness Inspection for Pseudo-screen

From the result of previous experiment, three patterns (EW, LD and PD) are used as stimuli for next experiment. In this experiment, we verified whether the connection of shots with different depth position will be recognized as a scene, from the view point of montage.

6.1 Stimuli

In each pattern, pseudo screen was located in the S3D space which has the convergence point at physical screen. The image of normal angle of shot was set as "shot 1" and following image of wide angle of shot was set as "shot 2", and these images are constituted as a scene. As for the position setting of pseudo-screen, shot 2 was set at convergence point (at the position of physical screen). Shot 1 is set at deeper position from physical screen. We verified whether viewers could feel that shot 2 came close to them. In addition, it is considered that the differences of distance between the image of wide angle of shot and the image of normal angle of shot would

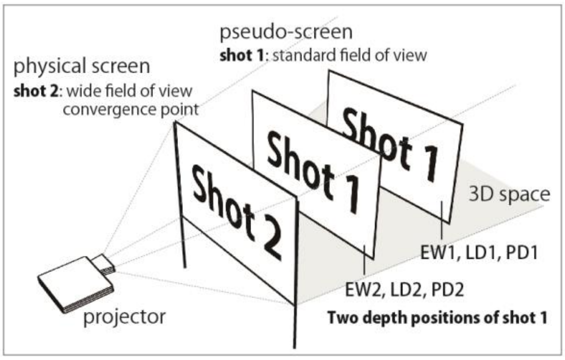


Fig.10 Concept of experiment 2

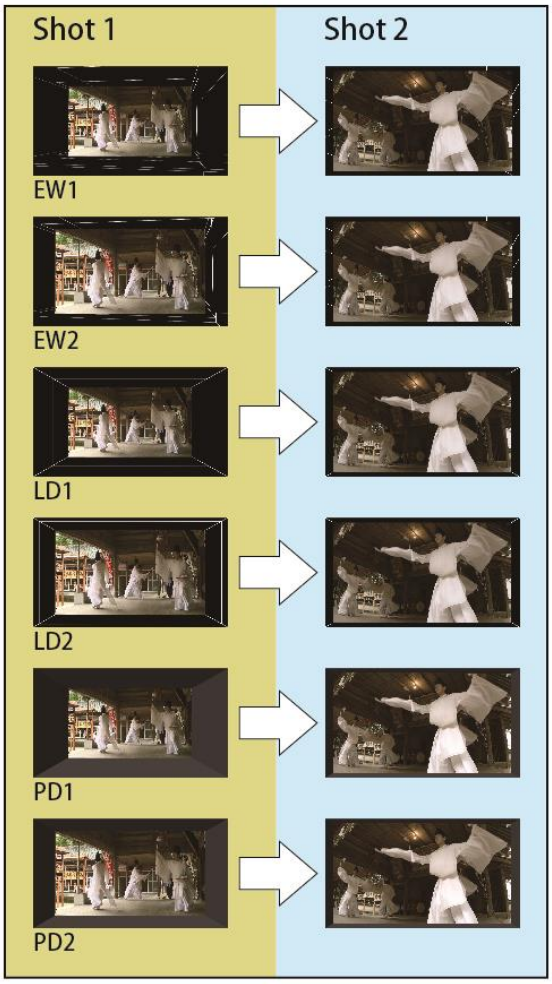


Fig.11 Visual stimuli for experiment 2

affect the impression of the shot, then two depth positions of shot 1 were set (Fig.10).

1. Depth position was set where visual size of pseudo-screen is 78 percent of the physical screen. **(EW1, LD1, PD1)**
2. Depth position was set where visual size of pseudo-screen is 64 percent of the physical screen. **(EW2, LD2, PD2)**

Thus, six scenes are created as visual stimuli for the experiment (Fig.11).

The procedure of the scene creation is similar to that of previous experiment.

Each visual stimulus has 20 seconds (shot 1: 10 seconds, shot 2: 10 seconds)

6.2 Experiment Procedure

20 students from Kyushu University participated the experiment (15 males, 5 females, all Japanese). All participants had experience to watch S3D movie. The experiment was carried out in the studio with condition of no illumination. The devices for the experiment are the same as previous.

Paired comparison method was used for this test. Combination all six stimuli including anteroposterior order are tested (30 pairs). For example, precedent stimulus EW1 (named as "A") and following PD2 (named as "B") were displayed, each stimuli was displayed 10 seconds with no sound. After paired ("A" and "B") stimuli were displayed, a white blank image is displayed twenty seconds, participants responded to five evaluation items (details will be described below) by five phases of evaluation standards. The evaluation paper was notation in Japanese (Fig.12). Two sessions were carried out including a break (two minutes) as one session with 15 pairs.

In the experiment, we inspected as to whether the viewing distance is changed and the scene is emphasized, by the combination of position change of pseudo-screen. Evaluation items are as follows.

"**visibility**": Visibility of the display method
"**close-up**": Sense of watching the screen in short range
"**montage**": Semantic combination of shots as montage
"**farness**": Recognition of distance
"**size**": Unity of subject's visual size

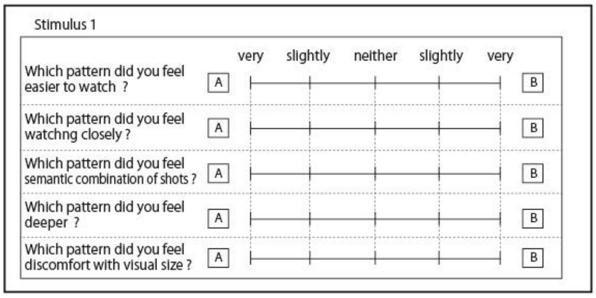


Fig.12 Evaluation sheet for experiment 2

6.3 Factor Analysis

Based on the collected data, two-way analysis of variance (factor 1: display pattern, factor 2: depth position of shot 1) and ranking by the evaluation point are performed. As for the analysis of variance table, in all evaluation items, interaction between factor 1 and factor 2 is not confirmed. Regarding "visibility" and "montage", there were no significant differences between the two factors (Table.3).

"close-up": In factor 1, there is significant difference between levels ($p < 0.05$).

"farness": In factor 2 there is significant difference between levels ($p < 0.05$).

"size": In factor 1, there is significant difference between levels ($p < 0.05$).

6.4 Discussion

As for "visibility", significant difference is not confirmed. Though pattern "LD" is highly evaluated. It is thought that the pattern that expressed perspective in a simple line is easy to look in the combination of the shots. Regarding "montage", there is no significant difference too, however, according to the oral interview, all participants did not have sense of incongruity of the combination of shot 1 and shot 2 as montage. It is considered that participants recognized the unity of shots by different screen size, as one scene. As for "close-up", PD1 was highly evaluated. It is considered that shot 1 is placed at longer depth position, participants felt that pseudo-screen approached more close to them. And about "farness", EW1, PD1 and LD1 were highly evaluated, and it is confirmed that pseudo-screen had an influence on sense of depth. As for "size", PD1 had the lowest evaluation, though it had the highest evaluation in "close-up". As a result, it is considered that sense of unity of visual size depends on the depth position and pattern. Moreover, regarding the general average, LD1 and LD2 was highly ranked (Table.4), then it is considered that the suitable visual pattern is the image with simple lines that viewers can recognize perspective at least.

closeup *p<0.05

Source	SS	df	F	
Factor 1	1.63	2	3.53111	*
Factor 2	0.00375	1	0.01624	
interaction	1.39	2	3.01119	

farness * p<0.05

Source	SS	df	F	
Factor 1	0.07583	2	0.13817	
Factor 2	1.17041	1	4.2651	*
interaction	0.00583	2	0.01062	

size * p<0.05

Source	SS	df	F	
Factor 1	1.21583	2	3.24318	*
Factor 2	0.66666	1	3.5566	
interaction	0.28583	2	0.76244	

Table.4 Ranking with average scores

rank	average	score	visibility	close_up	montage	farness	size
1	LD2	0.66	LD2	PD1	LD2	EW1	LD2
2	LD1	0.64	LD1	LD1	PD2	LD1	PD2
3	PD1	0.5825	PD2	LD2	PD1	PD1	LD1
4	PD2	0.55	PD1	EW1	LD1	EW2	EW2
5	EW1	0.525	EW2	EW2	EW2	LD2	EW1
6	EW2	0.49	EW1	PD2	EW1	PD2	PD1

Table.3 Analysis of variance table
("close-up", "farness" and "size")

7. Conclusion

In this research, we examined the displaying method that Pseudo-screen is placed in the Stereoscopic 3D (S3D) space. Pseudo-screen induce changing the viewing distance artificially. And we verified whether the connection of shots with different depth position will be recognized as a scene.

It is considered that the suitable visual pattern for the display method is the image with simple lines to let viewers focus on pseudo-screen. And it is considered that the change of the position of the pseudo-screen did not have an effect on the connection of the shot as montage. Thus, the display method is confirmed to be effective in S3D contents production. However, it is needed to be inspected about the depth position of pseudo-screen affecting the unity of sense of subject's visual size.

We focused on the depth position of pseudo-screen, with no aspect ratio modified. As for the cause that felt sense of incongruity to the visual size of the subject, it is considered that some participants felt that pseudo-screen changed its size simply, sense of depth was not recognized. To solve this issue, it is necessary to inspect it about the effect by the change of the aspect ratio of pseudo-screen. A shape of the screen changes according with angle of field. Pseudo-screen with image of wide angle of view is modified more widely. We consider that dynamism of the image can be emphasized by enlarging the horizontal field of vision, minimizing the change of the depth position.

As a proceeding in future, we will inspect the relationship between the depth position and the shape of pseudo-screen that effects the emphasis of scene.

References

[1] Takashi Kawai, Hiroyuki Morikawa, Keiji Ohta and Nobuaki Abe, *3D Rittai Eizo Hyougen no Kiso*, Ohmsha, 2010.

[2] "Buzen no Iwato Kagura" Henshu Iinkai, *Buzen no Kagura 2009*, Kubote Shiryoukan, 2009.

[3] Mineo Bitor, Noriaki Kumata, Hirokazu Yamanoue, Jun Tanada, Tadashi Hasegawa, Koichi Katori, *A Study for Switching and Cameras arrangement on Stereospic program Production*, ITE Technical Report, Vol. 22, No. 17, pp. 25-30, 1998.

[4] Takashi Kawai, *Jisedai Media Creator Nyumon 1 - Rittai Eizo Hyougen*, Cutt system, 2003.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 24603016.

Received April 6, 2014; Accepted June 15, 2014

Contextual Teaching and Learning Using a Card Game Interface

Teles, Sylker
Kyushu University
sylker@kyudai.jp

Tomimatsu, Kiyoshi
Kyushu University
tomimatsu@design.kyushu-u.ac.jp

Abstract

With bases well established in the constructivism of Vygotsky, Piaget and Dewey, Contextual Teaching and Learning is about to help students not only to memorize contents but also to construct the relation between subjects in order to define their meaning. Contextual Teaching and Learning represents an important field of study and analysis that can be widely applied to classrooms and new media but lack of information and research, especially in Video Games and virtual technologies. This paper proposes a Card Game based User Interface that reduces the player's learning curve of the gameplay and help the game content to be assimilated during the player's progress. Such interface can be applied to educational and non-educational Video Games always providing context, i.e. meaningful content, trough text card's design. For research purposes, a contextual card deck based on Art History was developed and tested against the same theme of traditional game interface with buttons derived from Massively Multiplayer Online Games abilities. The provided results shows that Card Games as Video Game's interface can help to simplify user's interaction while giving the player subjective information and content. By playing a prototype with the proposed interfaces, users were presented with questions about Art History, which answers were in the game's interface. User and community driven design method was used in the prototype development. As Virtual Contextual Teaching and Learning is a concept yet to be well explored and defined, the present research aims to contribute with the notion of human's cognition development in simulated environments.

Keywords: Card Games, Video Games, Contextual Learning, Education.

1 Introduction

Today's world is a complex mix of realities. While the society is changing into a more connected global community, educators around the world discuss how the school and education can follow such lighting speed evolution.

The research here described proposes a combination of *Trading Card Games* and *Contextual Teaching and Learning* (CTL) principles to develop a better game interface, i.e. a better interaction between the system and the user, more fit to transmit information and promote education. It was based on the hypothesis that Playing Cards can carry much more information than buttons and icons and are more visual appealing.

In order to test the hypothesis, two game interfaces were created and tested: a Trading Card-based Interface and an Icon-based Interface. These interfaces were applied in a prototype specifically developed for testing purposes. Commercial Trading Card Games and Massively Multiplayer Online Games (MMOG) were analyzed for the design and development process of each type of interface.

For the prototype, user and community directed design method came in hand to create and develop the game context, essential in CTL process. 10 cards based on Art History were created and a similar interface was designed just with icons, as user test control element. 20 Brazilian users from two active online communities tested the prototype during 3 days. Each group of 10 people tested one of the proposed interfaces.

2 About Card Games

To better understand the developed interfaces, the Card Games will be here briefly described. Despite its obscure origin, it is possible to trace the European early card game from the fourteenth century onwards, probably coming from the East. Theories vary from Asian tablets and discs coming from chess to the further introduction into Europe by the Saracens through Spain. Benham [1] suggests however, that the European Playing Cards are distinct and original in its conception and as the timeline of Oriental games is doubtful, there's a good chance of these games' been created at same time in different cultures.

Independent of the origin, Playing Cards became very popular in a short period of time and many different games and decks have emerged since. The early Tarot and the well spread four-suit European pack have inspired different types of Playing Cards in the late 20th Century, mixing strategic gameplay with collecting activity. The Trading Card Games were created. Trading Cards appeared a century earlier as cigarette and baseball cards.

The World War II and its consequent lack of paper has stopped the culture of Trading Cards temporary and ended the cigarettes cards for good. Sometime later, however, the Trading Cards went back as Tea Cards and Chewing Gum Cards. Until then, they were just Trading Cards. By adding game rules to them, the Trading Card Games were created. *Magic: The Gathering* is considered the conceptual father of all Trading Card Games [2, 3].

Besides *Magic: The Gathering* (M:TG), *Yu-Gi-Oh* cards were used as source for product analysis. Differently from M:TG, *Yu-Gi-Oh* is an example of media mix product similar to other Japanese famous card games as *Pokémon* and *Digimon* [4].
Following the path of many other media, the Trading Card Games were inserted into the virtual world, especially on the Internet. The number of online card games is beyond count; many of them are available in Social Network Services and mobile devices as well. Even M:TG and *Yu-Gi-Oh* have their virtual versions.
Unfortunately, just a few games had ever used cards to communicate their functions to players and receive input from them. Most of these games are from past two or three game console generations and have no glory to boast about. Still, some were good enough to be mentioned as *Lost Kingdom* series, *Kingdom Hearts: Chains of Memories*, *Phantasy Star Online Episode III: C.A.R.D Revolution* and *Baten Kaitos* I and II; and more recently *Metal Gear Acid* series.
Two types of gameplay could be identified amongst these games: *real-time battle system* and *strategic turn-based tactics*. In the case of *Baten Kaitos* series, the strategic element is not present leaving just the turn-based battle system. *Lost Kingdom* games have real-time gameplay, i.e. the character can move around make actions while the fight takes place giving more freedom to the player. By the other hand, *Phantasy Star Online Episode III: C.A.R.D Revolution* gameplay relies in a complex strategic combat involving many characters at same time. *Metal Gear Acid* series follows the same pattern and even the character movements are card-controlled.

3 About Contextual Teaching and Learning

As the hypothesis assumes that the proposed interfaces can promote education by context, the CTL concept will be explained from this point in parallel with education theory in Video Games.
Games, as a ludic activity, have a lot in common to CTL. The very nature of play is directly related to the learning process [5]. The learning process comes embedded in Video Games even though game designers don't have the intention to author an educational content [6].
Its built-in learning feature however, it is not enough to state Video Games as CTL valid tools. A comparison between Video Games' characteristics and CTL principles may enlighten the question. According to Sato [7], there are seven principles in context education approach: Purpose, Building, Application, Problem Solving, Teamwork, Discovery and Connection. Also, Crawford [8], states a set of five procedures called REACT: Relating, Experiencing, Applying, Cooperating and Transferring.
Starting from the Sato's list, *purpose* is about the meaning of doing something. The act of play is defined by Huizinga [5] as a "significant function". Games are meaningful and the meaning of play is of contextual nature, for the play act is not only about the game itself but the sum of players' experiences [9]. That is what Crawford calls *relating* in the second list above, i.e. "learning in context of one's life experiences or preexisting knowledge". The meaning and knowledge carried by an individual are therefore, outcome of one's own experiences. Without experiences the individual is empty. Experiencing can be considered the most important part of the process of learning by context because it involves the prime principle of learning by doing [10]. Therefore and furthermore, without experience or prior knowledge, CTL approach is not possible [8].
Teamwork and *discovery* from Sato's list represents what Crawford pointed as *experiencing*. It is not only the main concept to be part of CTL, but is also the core concept behind Video Games. Everything an individual does in life is experience of course, but interactive experiences are of more value to the learning matter [11].
Video Games are all about experiencing and immersion. The very act of play implies interactivity [9]. It contains the problem solving aspect of it [12] and is full of exploration, discovery and invention.
Applying can be found in both lists, and is subsequent action of *experiencing*. Crawford defines it as "learning by putting the concepts to use" [8]. In a very simple way, applying is experience again but now knowing what to do. The success of application can be motivating or frustrating. It cannot be too easy or too difficult to accomplish.
Next there is the *cooperating* aspect of Crawford's REACT. It can be also related with *teamwork* and *connection* in Sato's approach. In fact, the cooperating or sharing is a present point in the experiencing and applying when it comes with work in groups. It's true that some Video Games have no cooperation, but especially after the Internet, almost every game has the capability to share the user's experience.
When it comes to cooperation, many games offer the multiplayer option online and offline. In that sense it may be that the MMOG represent the most cooperative of games, with its thousand of players sharing the same simulacrum and experiences in a synchronous and persistent environment [12].
In order to advance in a MMOG, the player must engage in a collaborative learning process. It means socialize and cooperate in order to learn [12]. Lone gamers can have a hard experience by trying to advance in a MMOG. Most of times players must form small groups to overcome challenging dungeons or bigger ones to accomplish difficult raids or even more organized associations as guilds which members can better cooperate and share [13].
Collaborative learning also comprehends *knowledge transfer*, the last element of Crawford's list. Among the four ways of knowledge conversion proposed by Nonaka in *A Dynamic Theory of Organizational Knowledge Creation* [12], *externalization* is the most suitable in the sense of what Crawford suggests. According to the authors, the experienced players tend to share their knowledge with beginners, and that is what externalization means and also what *transferring* is about in CTL context. Thus, MMOG offer the best environment to apply the CTL concepts and promote a technology-mediated collaborative learning.
It is now possible to convert the found CTL principles and the Trading Card Games concepts into a Game User Interface that can be tested with users.

4 Interface Development

A simple method involving design needs and requirements [14, 15] was applied in the card development and heuristic evaluation [16, 17, 18] for the prototyping process. If something Playing Cards are complex information holders full of signs and playful content. In that sense, a user-centered approach based on cognitive and communicative aspect of interaction [19, 20] was used to fulfill the product requirements in combination with a community-centered design approach [21].

Trading Card Games are very social activities. Video Games in general are becoming more social nowadays. As discussed above, MMOG can be turned into a great tool for CTL just because its social features. Considering all these social factors, a community-centered design approach was really useful. Furthermore, it helped to define the target audience. By choosing a community in the first place, its users naturally composed the test group that the interface was designed towards.

Following that direction, two Brazilian communities with common interest were chosen. They are organized in the Social Network Service *Facebook* under the names *Nerd Power* and *SkyNerd* with 3.660 and 1157 users respectively in the time of writing.

The communities were chosen not only for its size, but also for the social learning element found in it. As the names suggests, they share common interest on *nerd* content, what varies from pop culture to quantum physics. The nerd community around the world has been bound to a negative stereotype for a long time [22], but the past two decades the nerds and geeks, or the *Third Culture* [23], has become popular and important to the media and industry.

Nerd communities usually share and discuss their subjects with great enthusiasm. It carries many constructivist elements associated with some relevant content, what promotes CTL. The communities here observed use different kinds of media to share and discuss content, including podcasts and video casts.

As mentioned, a game prototype was developed in a community-centered design. The game itself is part of a bigger research about CTL and Video Games. It is a Massively Multiplayer Online Game which narrative comes from many distinct community's resources. A Video Game needs an interface as any other tool, and in this particular case, the Trading Card Game proposition was applied and evaluated.

The game called temporarily Skynerd Protocol [24] is set in a apocalyptical Earth, when natural resources are scarce and knowledge of the human race of late is lost but for some artifacts to found during the game progress, artifacts that represent human History, Science and Art. The game's context is also well placed over environmental issues, as climate changes and preservation of natural resources. On top of all that, social, politic and economic discussions are contextualized in a world that has suffered with war and must now survive an atomic winter.

The context the game is based on can give room to many and more contextual learning. In fact, the game is been developed to be a collaborative platform for the *Nerd Power* community and educators alike. That is no matter for the research however, but all that contextualization may be well used alongside with a proper interface.

For the proposed experiment, a simple game mechanics was designed following the *real-time battle system* observed in some games earlier during the research. The game's goal is to find five hidden chests in which famous original paintings are stored. To open each chest the player needs to use a card corresponding to the painting artist. A deck of 10 cards was created (Fig. 1) and 4 cards are drawn each time for the player to choose.

Figure 1 Card Deck of Famous Artists

The cards hold information about the artists as year of birth and death (if applicable), brief bio and a quote, designed accordingly with the Card Games previously analyzed. Because the target audience is Brazilian, the information is written in Portuguese. Following the community-based approach, the cards were designed based on comic books, a common media among the target community.

The second interface was based on the traditional MMOG, such as *World of Warcraft* and *Star Wars: The Old Republic*. Those games make use of common icons to represent character abilities or items and labels which bear the necessary info to be viewed when the user roll the mouse over it. Again, the interface comes with the artists shown in the last figure and the info only appears in the mouse over state of the icon. The button or icon-based interface does not require the same gameplay as the Trading Cards; therefore, all icons are visible and available for the user all the time as seen in the Figure 2. The Figure 3 shows the Trading Card version of the interface.

Figure 2 Icon-based Interface

Figure 3 Trading Card-based Interface

Once inside the game, the user must answer 5 questions about Art History regarding the paintings and artists in the game. After completing the game the user is prompted with the same questionnaire again. The difference between the answers may give some directions on which interface is better for CTL and if the learning process is in fact occurring.

5 Results

The game prototype was tested with 20 volunteers from the mentioned communities, 17 male, 18 students, all Brazilian from 10 different cities and 9 states during the period between 7th and 9th of January of 2014. The prototype was programmed to run just once and guide the user through the first set of questions about Art History just after filling a form with personal data. To the user was given plenty of time to explore the simulated environment and find 5 pieces of lost famous paintings. Half the players used the Trading Card Interface and the other half the common MMOG abilities buttons. At the end of the game, the users were prompted with the same questions already answered but with a new question in place of the last one.

Also, the level of gaming experience was questioned with most of participants stated as regular (55%) or advanced players (40%). Just 5% have said they're beginners. It was important to establish whether the game mechanics or the interface complexity could interfere in the final results. Despite of a well balanced distribution, 80% of the players affirmed to play more than two hours per week, 60% more than 5 hours. It was considered enough for the simple game mechanics presented and to follow the interface instructions. The game playtime feedback (Figure 5) also provided at the end of the test showed that Card UI players finished the game in an average time of 6'18'' against 38'52'' of Icon UI users.

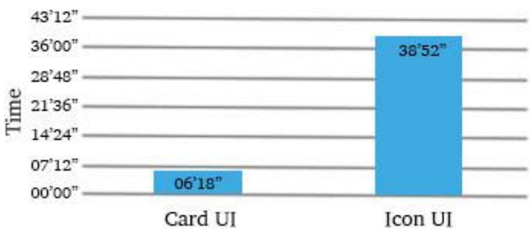


Figure 5 Total Playtime

The result demonstrates that Card-based UI promotes a better comprehension of the game mechanics and help players to end tasks quicker, also confirmed by applying a t-Test method.

The difference between the five questions presented before the gameplay and the five showed at the end are represented in the Figure 6. Although 40% of the players showed no evolution, in general, the after game questionnaire proved to have the best performance, including the only time when users got the top score. Besides, the second score was never lower than the first, been equal or higher. The average test score for Card UI players was 2.5 and 3 for before and after respectively. The same result for Icon UI users was 1.9 and 2.9. We can see here that Card Game Interface held better results in both questionnaires, but it does not mean it is the best for the learning process, at least in regard of the first test. Also, the learning curve was better for the Icon UI players.

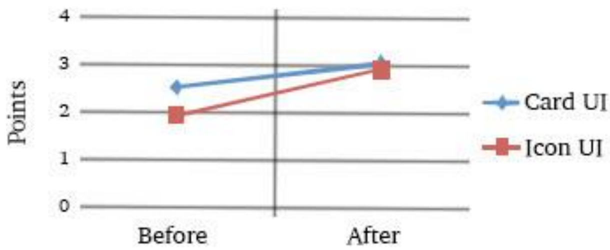


Figure 6 Game Score Comparison

Regarding the game total score, the graphic in the Figure 7 demonstrates a better performance for users using the Card Game UI. For each correct button or card used, 100 points were given to the player. Each wrong choice resulted in minus 50 points. The average score for them was 155 points against 100 for Icon users. Cards UI players were not only faster but got a better general score in the game too, also confirmed by applying a t-Test method.

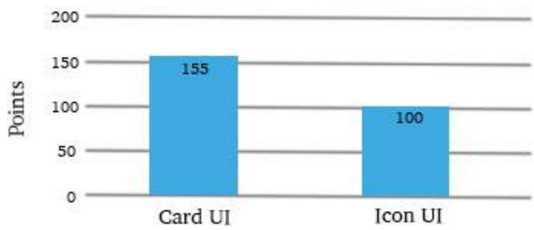


Figure 7 Game Interface Comparison

6 Conclusions

Of course there is more to do than what was done in this brief research. The direction is however defined. It is possible to state that Card-based Interfaces are helpful and can be used to build a faster and simple gameplay with a good learning process alongside. It's a better solution for CTL games.

Why exactly the cards offer a better choice to CTL is still unclear. Supposedly the cards carry more information visible to the player in real-time while buttons show only icons and numbers, with the need of a mouse-over to access the details.

The hypothesis can only be confirmed with users evaluation in future works.

The approach of two questionnaires, before and after game, showed a flaw. It was observed that most players tried to repeat the answers from one test to another. The observation was possible by adding a slightly different question in the second test. A better approach is to apply the tests with players and non-players and compare the results. It should be done with an expanded prototype that will have an improved gameplay, promoting user's collaboration, what lacked in the current application, preventing to prove Sato's [7] principles of Teamwork and Connection and Crawford's [5] Cooperating and Transferring.

About CTL's principles, the current prototype was developed to have all of them but the multiplayer factor, not available for technical reasons. It will be included in the next, more detailed, research.

The present work is relevant to point such directions. Learning and cognition process associated with Video Games are very complex and require a step-by-step approach. The results here presented are just the first step towards a complete interface proposition. The card-based interface here used and proposed worked well, but there is always room for improvement. In fact, is possible to conclude that the room is wide. User tests can help to build a proper card interface for CTL, although the one applied here was a good start indeed.

The complexity of information and educational content can also be improved. The quality of information and evaluation methodology will be pursued in order to allow more accurate and better results to analyze.

Despite those questions, some answers were also found. First of all, Cards UI promoted a better, faster and intuitive gameplay what resulted in good game scores and information absorption. Icon UI proved to be less clear and complicated to use, even though it is the most common interface in MMOG and most players were advanced game users.

Icon UI players, however, learned more during the process. While is known that it is not related to the interface because the before-game test was not influenced by UI, it's matter that needs attention in the future tests. The answer can help build better card elements and improve the learning curve for card-based interfaces as well.

Card Game-based Interfaces can be used in CTL games, but its format is yet to be improved.

7 Next Steps for the Research

This is a work in progress as any other research and moreover, just half of the interface development for all that is known about UI design. Much and more are yet to be done, starting with the user tests and interaction evaluation and most important, user socialization and collaboration. In the brief research here described, the multiplayer capabilities of the learning process were not observed and must be the next natural step.

In the technical side of the work, the client and server communication will be added to the prototype in order to turn it in a proper MMOG. It was found a new hypothesis that such kind of game can promote better the CTL than other genres. Also, the gameplay will suffer a drastic change by addition of a chain quest system, which will contain the educational content associated with problem solving situations.

Furthermore, questions have been raised. Why many users did not increase their score between sections and why some had even decreased it? The exact reason why the Card Game Interface was better is not clear but to the concepts previously discussed. The nature of it must be better investigated and the interface improved for the next researches. Chance is that the answer is related to the first two questions and may present a different path to follow as well.

For now is possible to assume that some level of contextual learning can be done through a Video Game with a proper interface. Whether the game context helped or not is another question for the near future. A proper ambient will be needed for that and a narrowed target audience as well. A controlled test inside a school's computer lab may serve the purpose.

It is also possible to state that the Trading Card Interface helped to improve the scores. Better investigation on this matter along with design improvements over its elements may enlighten the hidden processes that allowed those results and point to the design and development guidelines for future games.

These results can point easily to some topics discussed before, as the principles proposed by Sato [7] and Crawford [5]. The prototype was developed with problem solving and discovery in mind, with a big world to be explored. It carried purpose, building and application but it must be admitted that it lacked the connection, cooperation and transferring principles, yet to be developed and tested.

The interface here proposed is not new in games. It is not even new in MMOG, although scarce used in history. The research has found however that it can help improve the information transferred to the user during gameplay and the future researches will follow the same direction toward better games that can also bear useful and educational data whilst entertain the player in its contextualization.

There's a lot of work ahead and here some important points became clear whilst some others appeared. The future research lays over the improvement of the prototype; the adequacy of method and approach; the application of user tests and evaluation; and a better way to contextualize educational content inside a game.

References

[1] Benham, W.G. (2013). Playing Cards: History of the Pack and Explanations of its Many Secrets. Mccutchen Press.

[2] Lenarcic, J., James M-S. (2005). Trading Card Games as a Social Learning Tool. Australian Journal of Emerging Technologies and Society. Vol. 3, No. 2, 58-70.

[3] David-Marshall, B., van Dreunen, J., Wang, M. (2010). Trading Card Game Industry: From the T to the C to the G. Retrieved December 12, 2013, from SuperData Research Website:
http://www.superdataresearch.com/content/uploads/2009/08/TCG2010.pdf.

[4] Mizuko, I. (2003). Technologies of the Childhood Imagination: Media Mixes, Hipersociality, and Recombinant Cultural Form. Items & Issues, Vol. 4, No. 4, 31-34.

[5] Huizinga, J. (1971). Homo Ludens: A Study of the Play-Element in Culture. Beacon Press.

[6] Gee, J. P. (2007). What Video Games Have to Teach Us About Learning and Literacy. 2nd Edition. Palgrave Macmillan.

[7] Sato, H. (2001). 学習の意味探求を重視した文脈学習理論：デイルパネル著『なぜこれを学ばなければならないの』を読んで [Contextual Learning With Emphasis on Meaning: The Dale’s Panel “Why do I have to learn this?”]. 公教育システム研究, 1:173-178.

[8] Crawford, M. (2001b). Teaching contextually: Research, rationale, and techniques for improving student motivation and achievement. Texas: CCI Publishing.

[9] Salen, K., Zimmermar, E. (2004). Rules of Play: Game Design Fundamentals. Massachusetts Institute of Technology.

[10] Crawford, M. (2001a). Contextual teaching and learning: Strategies for creating constructivist classrooms. Connections 11(6), 1-2, 6. Retrieved June 11, 2013, from http://www.cord.org//uploadedfiles/Vol11No6.pdf.

[11] Chang, E. (2006). Interactive Experiences and Contextual Learning in Museums. Studies in Art Education, 47(2) 170-186.

[12] Kong, J.SL., Kwok, R. CW. (2009). MMOG Game-Based Coolaborative Learning: An Exploratory Study and its Research Potential. Procedings of Pacific Asia Conference on Information System.

[13] Nardi, B., Harris, J. (2006). Strangers and friends: Collaborative play in World of Warcraft. Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work, 149 – 158.

[14] Hartson, R., Pyla, P. S. (2012). The UX Book: Process and Guidelines for Ensuring a Quality User Experience. Morgan Kaufmann.

[15] Stone, D., Jarrett, C., Woodroffe, M., Minocha, S. (2005). User Interface Design and Evaluation (Interactive Technologies). Morgan Kaufmann.

[16] Nielsen, J., Molich, R. (1990). Heuristic Evaluation of User Interfaces. CHI '90 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Retrieved January 14, 2013, from ACM Digital Library:

[17] Pinelle, D., Wong, N., Stach, T. (2008). Heuristic evaluation for games: usability principles for Video Game design. CHI '08 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Retrieved 23 March, 2013, from ACM Digital Library: http://dl.acm.org/citation.cfm?id=1357282.

[18] Rogers, Y., Sharp, H., Preece, J. (2011). Interaction Design: Beyond Human – Computer Interaction. John Wiley & Sons, Inc. New York.

[19] Johnson, J. (2010). Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Rules. Morgan Kaufmann.

[20] McKay, E. N. (2013). UI is Communication: How to Design Intuitive, User Centered Interfaces by Focusing on Effective Communication. Morgan Kaufmann.

http://dl.acm.org/citation.cfm?id=97281.

[21] Preece, J. (2000). Online Communities: Designing Usability and Supporting Sociability. John Wiley & Sons, Inc. New York.

[22] Kendall, L. (2001). White and nerdy: computers, race, and the nerd stereotype. The Journal of Popular Culture, 44(3) 505-524.

[23] Kelly, K. (1998) The Third Culture: Essays on Science and Society. Science 279(5353) 992-993.

[24] Da Silva, S. T., Tomimatsu, K. (2013). Game Prototyping with Community-driven Narrative. Proceedings of The 2nd IEEE Global Conference on Consumer Electronics (GCCE). 376-378.

Received April 6, 2014; Accepted June 15, 2014

International Journal of Asia Digital Art & Design

Gocen: Appropriating Simplified Handwritten Notation as a Musical Interface

Tetsuaki Baba
Tokyo Metropolitan University
baba@tmu.ac.jp

Yuya Kikukawa
Tokyo Metropolitan University
kikukawayuya@gmail.com

Kumiko Kushiyama
Tokyo Metropolitan University
kushi@tmu.ac.jp

Abstract

Not only in childhood but also adulthood, we need some training to read music scores, which sometimes make music hard to learn and enjoy. In this article, we shall propose the system that enables users to play their handwritten musical notations by our musical interface. Since 1960's, Optical Music Recognition (OMR) has become matured in the field of printed score. In recent, some products were released on a market that uses OMR for music composition and playing. However, few researches on handwritten notations have been done, as well as on interactive system for OMR. We combined notating with performing in order to make the music more intuitive for users and give aids to learn music to users. We set 6 design criteria for evaluation of our musical interface. Through users tests and user observations, we shall report our system has high accurate image recognition for handwritten notation and is success to help users to create music along with our design criteria.

Keywords: handwritten notation, optical music recognition, optical character recognition, interaction design

1 Introduction

We are interested in developing tools that support the enjoyable education and performing process for music. Since so many digital musical instruments have been developed, the interface of electronic musical instrument has been evolved new types, such as gestural, wearable, computer vision, etc.

From the imitation of the classic musical instrument to a totally new one, these kinds of digital instruments provide new possibility for performance and education of music. To learn and performing some piece of music, we generally use staff based score that consists of five horizontal lines and notes. The staff originated from musically annotated text, though the Gregorian Chants around the 12th to 13th centuries. It is a basic and important literacy that not only professional musicians but also beginners can use staff based score for learning, playing and composing. Professional musicians have literacy to read and write the staff. But other than the classical music, some users and semiprofessional users often cannot read and the write score.

Not only in childhood but also adulthood, we need some training to read music scores, which sometimes make music hard to learn and enjoy. In this article, we shall propose the system that enables users to play their handwritten musical notations by our musical interface. By combining notation with performance, our system makes music more intuitive and accessible to enjoy and learning music. Concretely, users can play by scanning their handwritten notation with a kind of scanning device. Users can do intuitive interaction: play, re-play and play backwards. Furthermore, users can make some articulations, such as pitch bending and a vibrato by moving scanning device during playing.

Figure 1 shows sketch image of our system. The user writes simple handwritten score on a paper, and then makes sound by tracing with the device. As mentioned next section, it is difficult to recognize and process handwritten score as an interactive system for the contemporary research field of Optical Music Recognition.

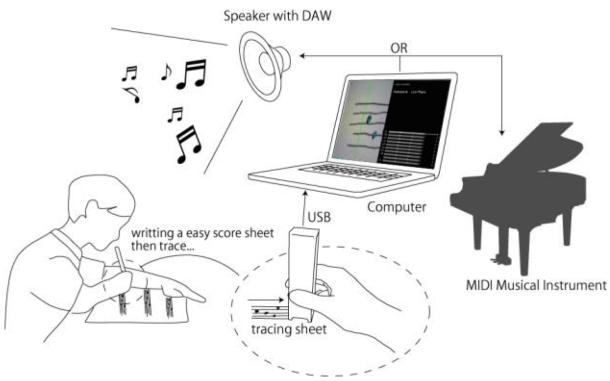


Figure 1 Abstract of our system

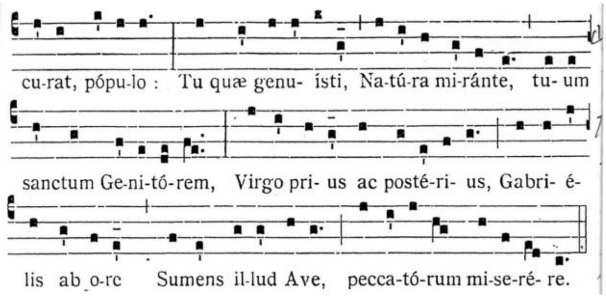


Figure 2 A sample score of Gregorian chant

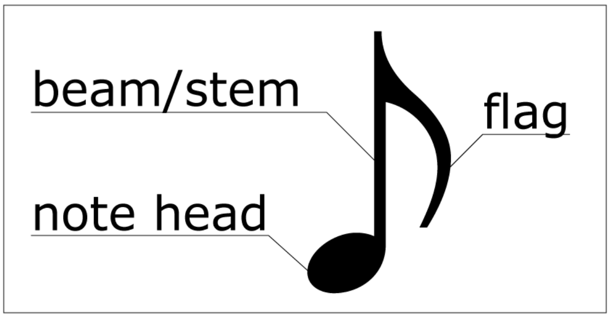


Figure 3 Composition of each part of musical note

In this study, we resolve handwritten recognition accuracy and speed of processing by simplify scores that consists of note head and five line and some accidentals. This kind of score is similar to Gregorian chant. Figure 2 shows sample score of Gregorian chant. Then we design interface and interaction for the system. Note head is a part of note that is shown figure 3. We call this filled part “note head”.

2 Related Work

2.1 Optical Music Recognition

Optical musical recognition for handwritten can be roughly classified into two kinds of input methods: online and offline input. Online input is a method that users can input stroke information via pointing devices: touch-pen, mouse, etc. On the other hand, offline input is a method to scan optical information via kind of paper.

In the case of online input, There are several researches for applying stroke information to character recognition system [2,8], Especially Forsbergs [8] is well known as the simplified stroke based interactive system. Whereas, We uses offline input for optical music recognition (OMR).
OMR for offline printed musical scores has been matured and active research field. Since Kassler [14] in 1963, Pruslin [6] in 1966, Prerau [22] in 1970 have published papers on OMR. Up to the present time, OMR research field have been still active. In Japan, Since 1980, Miyao[19], Ohteru[20], Matsushima [18] have reported on OMR. Although many researchers are trying to enhance recognition accuracy, it is still difficult to recognize all kinds of music scores perfectly because of diversity of musical notation and symbols. On the other hand, some techniques of OMR has been applied to commercial applications and OMR libraries for developers: Audiveris [3], OpenOMR [5] and Gamera Framework [7] has been distributed.

As an example for interactive OMR, Kawai has released OMR smartphone application [15], which enables users to play music from musical scores easily. Kawai has also released “PDF Musician” [16] that users can play musical scores by touching and tracing. It provides users with familiar and intuitive OMR interactive system. This system uses text PDF that score information is implemented in. It could not realize musical scores precisely and instantaneously unless users prepare a PDF file that includes text PDF. Yamamoto and et al. [25] proposed musical interface system named ”on-Note”. They use physical marker less musical scores to play music intuitively. The notes on the score are captured by a camera and are processed by the system that retrieves the music from a score database. In addition, the system can do a real-time recognition of the paper’s position and the rotation. Thus, by physically moving and connecting the musical scores we can play music intuitively.

These kinds of researches are similar to ours on the view that users can directly use musical scores for playing music. But our methods are different from the view that users can use own handwritten score and interactive techniques for playing.

“The Music Wand” by Hoerter [12] is most similar to ours. A user can play music by using printed scores for playing with hand-held device. They reported recognition of note position is about 70 % and accidentals is less than note position. In our first prototype, we have used hand-held scanning device to read music score and its recognition accuracy is same as Hoerters. Furthermore, Their system needs a light box to get binalized image and a kind of ruler to support uses interaction, which puts an obstacle for usability.

Many researches are focused on printed scores, but some researches about OMR focused on handwritten score has been also reported since 1970s. Bulis showed that computerized recognition of musical notes can be accomplished by a relatively of comparing horizontal and vertical histograms of symbols. Yadid-Pcht and et al. reported recognition of handwritten musical notes, based on neural network model. And they showed average 80 % recognition rate for note recognition. In the case of interactive system for OMR, we have to develop a solution against recognition rate and responsibility.

2.2 Tangible UI & Media Art

In the field of media art or tangible UI, many researchers reported kinds of new musical instruments. Especially several researches or works are focused on enhancing comprehensibility for music performing and learning as a motif of music score.

Bottelo proposed Tangible user interface for music learning to help children learning the music theory. He developed the software system that enables users to create music by putting objects labeled optical marker on the table [4]. A user can not only put objects but also rotate or extend objects to make change pitch or length of note.

Noteput by Jonas Friedemann Heuer [11] is an interactive music table based on five line staff, which combines all three senses of hearing sight and touch to make learning the classical notation of music for children and pupils easier and more interesting. This kind of timeline based musical interfaces have already well known by Golan Levin [17], Toshio Iwai [23] and et al.

Drawn by Liberman [27] presents a whimsical scenario in which painted ink forms appear to come to life, rising off the page and interacting with the very hands that drew them. Inspired by early filmic lightning sketches, in which stopmotion animation techniques were used to create the illusion of drawings escaping the page, drawn presents a modern update: custom-developed software alters a video signal in real time, creating a seamless, organic and even magical world of spontaneous and improvised performance of hand and ink. Calligraphy is one of art form and it is the supreme art form in China. So we could not separate writing and paper. This kind of primitive and intrinsic interaction between human and paper is important for our creative activity.
Tsandilas and et al. [26] focus on the creative use of paper in the music composition process and proposed Musink to provide composers with a smooth transition between paper drawings and OpenMusic as a flexible music composition tool by using the Anoto pen. In another project, they insisted that composers can use paper effectively on the first stage of composing [9].
Our goal is to develop music performing and learning system that do not use mouse, keyboard and touch panel on a computer by using handheld scanning device and designing user interaction between users and paper.

3 Gocen

3.1 Design criteria

Poupyrev and et al. [22] mentioned design issues for musical controller for good design of musical interface on workshops, chi2001. Therefore we also same criteria for evaluating our interface. Furthermore we adds one more criteria "uniqueness" to theirs, which means single interface must control all interactions. We aim to develop a system that do not require the legacy and generic interfaces such as pc keyboard, mouse, touch pen and touch pad.

- Usability and comprehensibility
- Expressiveness
- Sensitivity and sophistication
- Aesthetics
- Hedonics
- Uniqueness

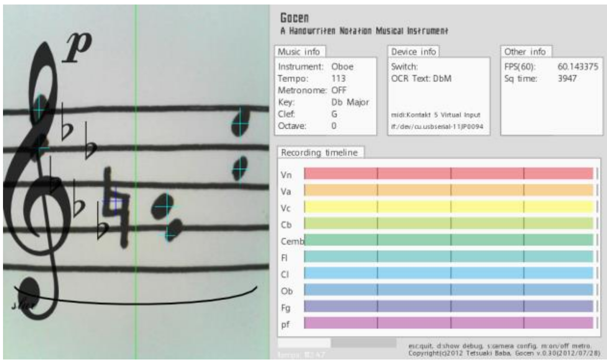


Figure 4 A screenshot of our software. Left captured image shows camera image on which is overlaid musical information. Right panel shows text information, timeline and more.

3.2 System abstract

Our system mainly consists of a scan device, computer and sound module. A user can play simple music by tracing notes with the scan device as shown in figure 1.

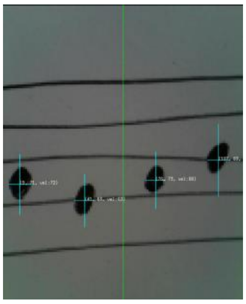


Figure 5 A screenshot showing recognized notes

The computer processes captured images by using OpenCV

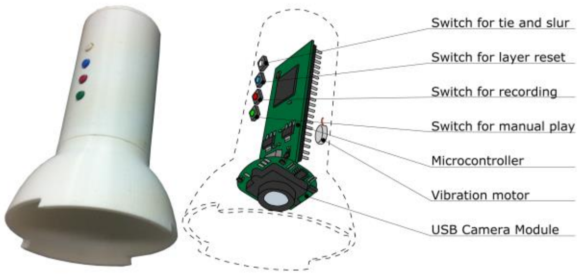
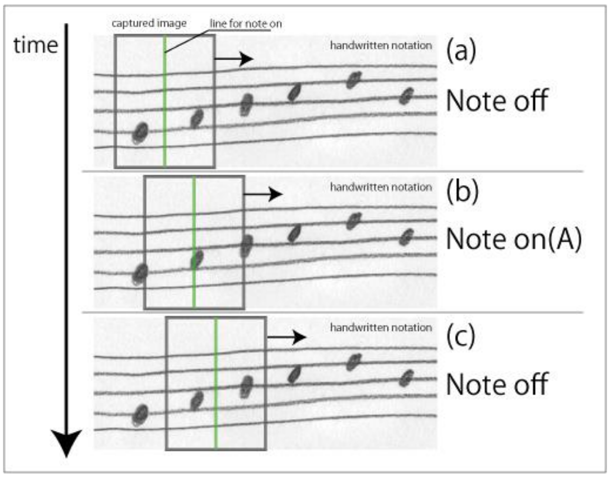


Figure 6 Gocen device consisting of USB camera, microcontroller, switches and vibration motor

and our algorithm at 30fps, then outputs sounds according to the data from the notations.

We do not need any special materials other than this system. A user can use normal white paper and his/her own pen. Our device is built with a USB camera, microcontroller, and vibration motor (see figure 5). The vibration motor is used for tactile feedback while the user is playing.



Our software is made on some libraries: openFrameworks[21], OpenCV, OpenGL, portmidi, Ocrad[1]. The system can select

Figure 7 Flow to make a note on/off

(a) (b) (c) (d) (e)

Figure 8 Example images of interaction. (a)Chord, (b)velocity, (c)octave, (d)key in A Major, (e)bass clef

Figure 9 Example of score of "Twinkle, twinkle, little star" that is shown to users.

Table 1 Result of the user test for pitch recognition.

trial	participants	total	correct(Ave.)	accuracy
1	21	42	40.9	97.4 %
2	9	42	42	100 %

any midi devices. On this stage, we use Kontakt Player[13] for MIDI device. Figure 4 shows screenshot of our software.

Our system can be roughly classified into two kinds of processing. One is OMR processing which is mainly control performing. Another is OCR processing which is a mainly control setting, such as key, key range, instruments and etc. We uses 3rd party library for OCR function. Therefore we do not refer to recognition accuracy about it.

3.3 Interaction

Our system is not only an OMR system but also performance system. We developed several musical interactions for this interface.

Note on/off
First of all, a user write 5-line staff and note head on a paper. Then a user holds the handheld device and put it on the paper. A scanned image will be shown on the screen. Figure 6 is a screen shot of a image that shows each position of note head as cross hairs. The vertical size of a cross hairs shows velocity of the note. Figure 4 shows a screenshot of application on PC screen. A user can confirm some information: selected instrument, options, timeline for recording and so on. While pushing green button on the device, a user can play each note head by moving left/right to right/left over a note head. Figure 7 shows flow to make a note. To place green bar on the screen upon a note, a user can make a sound.

As a result, these kinds of operation enable uses to ignore the information of musical note length. No matter how a note flag shows sound length. A user can move the scan position with the handheld device by his/her hand manually. That's why we simplified the staff-based musical notation.

Chord
Our system can recognize a chord up to 32 notes at the same time. Figure 8(a) shows example notes of a chord.

Velocity
It is poplar to use kind of symbols: such as pp, mp, mf, ff and etc. for instructing velocity of a note. For example, a user can change velocity into mezzo piano to put the device on a word mp that a user wrote on a paper. Whereas, a user can also control velocity with the vertical size of a note. Larger note makes louder sound. Smaller note makes small sound. Figure 8(b) shows that our system changes velocity depending on the vertical size of a note.

Pitch Bending
In the case of kind of string instrument, such as violin, musician expresses articulations: vibrato, turn and etc. techniques. In our system, a user can use pitch bending by moving the device up or down after making a note on. The volume of pitch bending is defined as $|Y_1 - Y_2|$. Y_1 is first vertical position of a note on. Y_2 is the vertical position of the phonetic note. In the case of $(Y_1 - Y_2 < 0)$, the pitch is increased.

Tonal Range
The default tonal range is set from C4 to A5. It can be extended by writing an additional line or octave mark. In our system, a user can select the tonal range to read some characters such as "+15", "+8", "0", "-8", "-15" with OCR function (see Figure 8(c)).

key
The default key is C major. To change the key of score, a user put the device on handwritten characters, "AM", "Am", "BM", and etc.

Key Transition
A user can make key transition by operating note on/off after

read with OCR function.

Changing a instrument
A user can change the instrument of sound by covering a text he/she wrote to indicate the name of instrument, such as pf(piano), bs(bass), gt(guitar), dr(drums), etc., with the device, by means of OCR.

Changing the kind of clef
A user can change the kind of clef to read "C.F" or "C.G" with the device. Figure 8(e) shows a status on clef F or bass clef.

Loop sequencer
A user can record sound events into timeline, while pressing recording button and make a loop like a sequencer. Each recorded note will be set in the quantized timeline.

Figure 10 Example of scores that users wrote.

Table 2 Result of the user test for accidental recognition.

	1st trial			2nd trial		
user	**sharp**	**flat**	**natural**	**sharp**	**flat**	**natural**
1	73 %	77 %	90 %	97 %	93 %	100 %
2	70 %	90 %	100 %	87 %	90 %	100 %
3	90 %	30 %	100 %	100 %	97 %	93 %
4	87 %	10 %	100 %	100 %	37 %	97 %

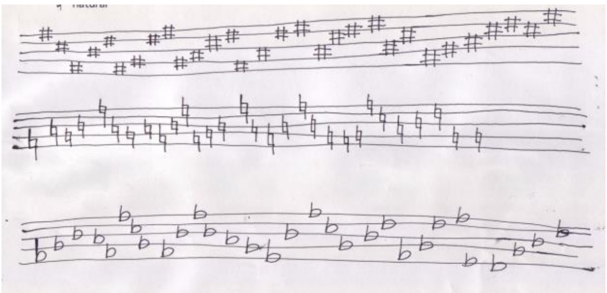


Figure 11 Example of scores that users wrote.

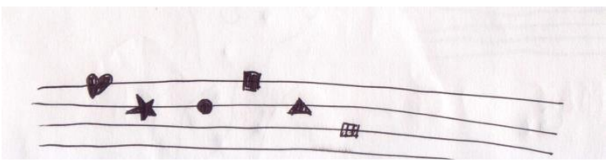


Figure 12 A user wrote no musical symbols

4 User Study

4.1 User test for recognition accuracy of pitch

We find how users can play a correct pitch of note they want by simple user test. We recruited 19 participants (9 male) from our university students who has experience of notation: "almost everyday": 0, "often": 3, "rarely": 8, "almost nothing":8, Age average is 21.9.

Before the experiment, we distributed a experiment paper and instruct about our system. After that, we let participants write a simple piece (twinkle twinkle little star) and play it. An example score had been already printed on experiment paper (see figure 10). It is made on DTP software because an example score might affect users' writing. Below is the procedure of experiment.

1. Instruction (5min)
2. The user writes and plays their handwritten notation.
3. If user finds some mistake about a pitch, the user modifies their score and plays it again.
4. If user finds some mistake, goes back to 3.
5. Free description questionnaire

We provided two kinds of pen. One is 0.7mm oil base ballpoint pen. Another is 0.5mm oil base ballpoint pen. Participants can chose the pen that they want. We also provided participants with general plain paper. All number of notes of the score is 42 which is shown figure 9.

10 users could play all notes correctly on the first trial. After modification, all other 9 users could play all notes correctly. Table 2 shows the result. Average recognition rate on the first trial is 97.4 % and 100 % on the second trial. Figure 11 shows sample sheet that users wrote on their paper. Users answered almost affirmative response about our system such as fun, interesting, this is an epochmaking, better to commercialize, great, If I had this system in childhood, I could practice piano harder, and etc. On the other hand, we got dissatisfaction response such as, it need precision for interaction, it makes better if I could use colorful pens, difficult at first.

4.2 User test for recognition accuracy of accidentals

We developed a contour-based CV algorithm with Support Vector Machine. Fig.13 shows an example flow of a sharp accidental symbol from an original image. We use edge detection features for machine learning algorithm.
We recruited 4 participants (2 male) from our university students. Before the experiment, we distributed a experiment paper and instruct about our system. After that, we let participants write 30 symbols for each symbols (sharp, flat, natural). An example score had been already printed on experiment paper (see figure 9). It is made on DTP software because an example score might affect users' writing. Below is the procedure of experiment.

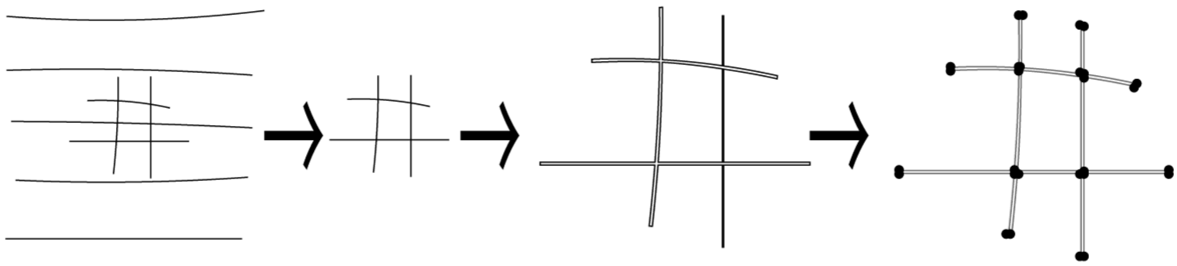


Figure 13 Our accidental symbol detection algorithm.

1. Instruction (5min)
2. 1st trial: The user writes 30 handwritten notations for each accidental symbol.
3. An experimenter gives him/her some advices about how to write recognizable.
4. 2nd trial: The user writes 30 handwritten notations for each accidental symbol.

On 1st trial, we showed participants printed accidental symbols. Figure 11 shows sample sheet that a participant wrote.

We had several exhibitions on domestic events and a international conference. Through exhibitions we have observing users carefully. Score notes that users wrote has 90 pages and over 300 users have experienced our system. A user draw symbols that is not musical symbol for his creative expression. Figure 12 is the image the user drew. We did not mention about other kind of symbols. This user showed this image to us, then suggested to implement to recognize these symbols. Many users spent more time for writing notations than OCR. Although we prepared many OCR settings, many users enjoyed writing notation and playing.

5 Discussions

Through the user study, we found many users can use easily our system whatever they has literacy to read/write music scores. On the user test, we found expected high recognition accuracy of pitch and accidentals, because of simplification of handwritten notation.

5.1 Validity of Design Criteria

We shall review the evaluation of our design criteria on the basis of the results of the user test and observation.

Usability and comprehensibility
Many users could play and enjoy with our system as soon as we told them how to play, because our common mental image of music notation help us to understand our system easily. Through exhibitions, we found that users who are around 6 to 60 years old enjoyed our system. Concerning with usability, especially design of our prototype device should be taken consideration. Several users could not get know how to hold our device, because of its no intuitive design. And we found tangled cables of the device interfered user interactions.

Expressiveness
As we mentioned before, our system has 10 interactions for playing. By using these, it enables users to play many kind of musical piece. Moreover we showed a possibility to extend our system by ensemble. But we have to verify the responsibility of our system carefully on future work. Musical score has more expressiveness than our system. For example, our system could not recognize two columns of 5-line staff. We have to continue develop the system constantly, but our system has achieved to a certain degree for expressiveness.

Sensitivity and sophistication
As a result that we developed specific OMR processing procedure, we implemented fast interaction with 30 fps. During the user test, one participant told us it is surprising that this system processes faster than he expected. At a recognition accuracy point of view, we showed our system has enough precision as interactive system.

Aesthetics
We have to continue to consider design about the shape of prototype carefully in future. We found some users often play music that they create with our system, it is caused that our system could provide users with creativity about music. Figure xx also shows that our system is success for providing users with contingency that is important human activity for art creation. In near future, we are going to create sophisticated performance to show how our system can make a attractive performing.

Hedonics
We could observe some children or users who can read/write score plays our system 30 minutes and more. We got comments about tactile feedback such as it feels good, feels like I playing

Uniqueness
We designed that a user can use all interactions with only one device. A result that we could not get any unsatisfied comment about from users shows a success for offering uniqueness of a device.

6 Conclusions

There are many researches about OMR. But few researches on interactive handwritten notation have been done. We proposed the system that enables a user to play their handwritten notation on a paper with our device by scanning. Our system can recognize handwritten note head and accidentals with high accuracy by simplifying musical notations. We could implemented not only accurate recognition system but also fast processing (30fps).The FPS except for capturing image is about 150 [fps]. It is possible to improve operability by improvement of computer and camera specification.

References

[1] Ocrad - the gnu ocr -. www.gnu.org/software/ocrad/ (last accessed, Jun, 2012).

[2] Anstice, J., Bell, T., Cockburn, A., and Setchell, M. *The design of a pen-based musical input system.* In Proceedings of the 6th Australian Conference on Computer-Human Interaction (OZCHI '96), OZCHI '96, IEEE Computer Society (Washington, DC, USA, 1996), 260–.

[3] Bitteur, H. Audiveris: *Open music scanner*, 2000 – 2012. http://audiveris.kenai.com/(last accessed, Jun, 2012).

[4] Bottello, L. *Tangible user interface for music learning*, 2010. http://cloverleafdesign.weebly.com/(last accessed, Jun, 2012).

[5] Desaedeleer, A. Openomr. http://sourceforge.net/projects/openomr/(last accessed, Jun, 2012).

[6] D.H.PRUSLIN. *Automatic Recognition of Sheet Music*. PhD thesis, Doctoral Thesis MIT Cambridge MA January 1967, 1967.

[7] Droettboom, M., Fujinaga, I., MacMillan, K., Chouhury, G. S., DiLauro, T., Patton, M., and Anderson, T. *Using the gamera framework for the recognition of cultural heritage materials.* In Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries, JCDL '02, ACM (New York, NY, USA, 2002), 11–17.

[8] Forsberg, A., Dieterich, M., and Zeleznik, R. *The music notepad*. In Proceedings of the 11th annual ACM symposium on User interface software and technology, UIST '98, ACM (New York, NY, USA, 1998), 203–210.

[9] Garcia, J., Tsandilas, T., Agon, C., and Mackay, W. *Interactive paper substrates to support musical creation*. In Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems, CHI '12, ACM (New York, NY, USA, 2012), 1825–1828.

[10] George, S. E. *Clustering on-line dynamically constructed handwritten music notation with the self-organising feature map*. In Proceedings of the 16th international conference on Developments in applied artificial intelligence, IEA/AIE'2003, Springer Springer Verlag Inc (2003), 93–103.

[11] Heuer, J. F. *noteput is an interactive music table with tangible notes, that combines all three senses of hearing, sight and touch to make learning the classical notation of music for children and pupils more easy and interesting*. http://www.jonasheuer.de/index.php/noteput/(last accessed, Jun, 2012).

[12] Hoerter, N., and Chatt, T. *The music wand: Real-time optical scanning of sheet music*, 2008. http://people.ece.cornell.edu/land/courses/ece4760/FinalProjects/s2008/tjc42_nah35/tjc42_nah35/index.html(last accessed, Jun, 2012).

[13] Instruments, N. Kontakt player. http://www.native-instruments.com/(last accessed,Jun, 2012).

[14] Kassler, M., and of Music, P. U. D. *An Essay Toward Specification of a Music-reading Machine..*. Princeton University, Department of Music, 1963.

[15] Kawai Musical Instruments Mfg. Co., L. Gakufu camera, 2011.

[16] Kawai Musical Instruments Mfg. Co., L. Pdf musician, 2012.

[17] Levin, G., and Lieberman, Z. *Sounds from shapes: audiovisual performance with hand silhouette contours in the manual input sessions.* In Proceedings of the 2005 conference on New interfaces for musical expression, NIME '05, National University of Singapore (Singapore, Singapore, 2005), 115–120.

[18] Matsushima, T., Kanamori, K., and Ohteru, S. *Automated recognition system of musical score (vision system of the wabot-2)*. Journal of the Robotics Society of Japan 3, 4 (1985), 354–361.

[19] Miyao, H. *Symbol recognition for printed piano scores based on the musical knowledge* (in japanese). The transactions of the Institute of Electronics 75, 11 (1992), 1848–1855.

[20] Ohtheru, S., Matsushima, T., and Kanamori, K. *Automatic recognition of printed music*(in japanese). journal of Acoustic Society of Japan 41, 6 (1985-06-01), 412–415.

[21] openFrameworks community. openframeworks. www.openframeworks.cc (last accessed, Jun, 2012).

[22] Poupyrev, I., Lyons, M. J., Fels, S., and Blaine (Bean), T. *New interfaces for musical expression*. In CHI '01 extended abstracts on Human factors in computing systems, CHI EA '01, ACM (New York, NY, USA, 2001), 491–492.

[23] Prerau, D. S. *Computer pattern recognition of printed music*. In Proceedings of the November 16-18, 1971, fall joint computer conference, AFIPS '71 (Fall), ACM (New York, NY, USA, 1971), 153–162.

[24] Toshio, I. Chess on the table, 1997.

[25] Tsandilas, T., Letondal, C., and Mackay, W. E. *Musink: composing music through augmented drawing.* In Proceedings of the 27th international conference on Human factors in computing systems, CHI '09, ACM (New York, NY, USA, 2009), 819–828.

[26] Yamamoto, Y., Uchiyama, H., and Kakehi, Y. *onnote: playing printed music scores as a musical instrument*. In Proceedings of the 24th annual ACM symposium on User interface software and technology, UIST '11, ACM (New York, NY, USA, 2011), 413–422.

[27] Zachary,L. Drawn. http://thesystemis.com/projects/drawn/ (last accessed, Jun, 2012).

International Journal of Asia Digital Art and Design Vol.18

No.02 Date of issue 2014/07/01

Asia Digital Art and Design Association

Chair

Kiyoshi, Tomimatsu

Chairperson, Academic Referee Committee

Hyunsuk, Kimi

Toshiya, Kuroiwa

Park, Seungho

Kuo-Kuang, Fan

Designed By Iwamoto, Yushi

Edited By Takase, Yuiko

Izumoto, Chika

Published By

Asia Digital Art and Design Association

c/o Faculty of Design Kyushu University

4-9-1, Shiobaru, Minami-ku, Fukuoka, 815-8540, Japan

http://adada.info adada.post@gmail.com