MYSQ : An Entertainment System Based on Content Creation Directly Linked to Communication

SATORU TOKUHISA, SUSUKE OKUBO
imgl/KMD (Keio Media Design)
KEN SUGURO, TAKU KOTABE,
Ututu Co., Ltd.
AND MASA INAKAGE
imgl/KMD (Keio Media Design)

This article is a description of MYSQ, an entertainment system that allows users to create original music video clips (ex:MTV) and share the resulting video with friends via cellular phones. The original movie is original because it is generated using video of the user, combined with video effects that use the user's dance movements as parameters. The MYSQ concept is communication through the process as well as the product of content creation. The MYSQ system can be characterized as: simple operation, multilayered interaction, collaboration, the conversion from player experience to content, and making experience transferable by converting it to media form. This article will then analyze the system with game flow, based on
Csikszentmihalyi's FLOW theory, to evaluate the system's entertainment value as a game, and the effectiveness of the characteristics above.
Categories and Subject Descriptors: J. .5 [**Computer Applications**]: Arts and Humanities
General Terms: Design, Human Factors, Theory
Additional Key Words and Phrases: Entertainment system, entertainment computing, creation, communication, collaboration, mobile movie, dance

1. INTRODUCTION

Since the inception of the concept of tangible bits [Ishii and Ullmer 1997], there have been numerous attempts to recreate physical sensations from computing. This trend can be seen in the field of entertainment computing. Specifically, a new style of game called Pergame tries to create social interaction as in traditional analogue games, instead of focusing solely on human to human or human to object interaction [Maerkurth et al. 2005; Benford et al. 2005]. There are studies being conducted in fields not related to games that apply physical and social interaction to entertainment systems built in order to create [Ryokai et al. 2004; Iwai and Nishibori 2005; Iwai and Sato 2005]. In these systems, the creation process itself creates entertainment value, which causes social interaction among the players. However, the social interactions that occur in these

Authors' addresses: S. Tokuhisa email: dangkang@sfc.keio.ac.jp; S. Okubo, email: sosuke@sfc.keio.ac.jp ; and M. Inakage, email: inakage@sfc.keio.ac.jp . imgl / KMD (Keio Media Design), S110, 5322, Endo, Fujisawa, Kanagawa, Japan; K. Suguro, email: magronbass@ututu.com; and T. Kotabe, email: taku@ututu.com. Ututu Co., Ltd., 201, 5-49-8, Jinguu-mae, Shibuya, Tokyo

systems are limited to direct communication, since they are collaboration s in which one player simply reacts to another player's actions.

This article introduces MYSQ (*My Style So Qute*!) [Ututu Co, Ltd. and MYSQ 2005], a system that combines game- and nongam-related social interactions and enables players to communicate regardless of time and physical location via the process and result of content creation. MYSQ uses a 2.5 cubic meter booth. Video effects are applied on a video of the player captured by the camera inside the booth using the parameters derived from the players' dance. The dance parameters control the video effects. Throughout this process, the player can create a fancy movie clip of his or her mall while enjoying the "creative" process. The resulting movie can be shared by sending it via cellular phones. Communication occurs between players who are playing at the same time, and also between players and potential players waiting outside the booth. After game play, the resulting video can be shared with friends, which causes communication with people who do not share the time or physical space with the player. MYSQ realizes these three different types of communication through the process and product of content creation.

MYSQ consists of the following five characteristics:

(1) Simple operation
(2) Multilayered interaction
(3) Collaboration
(4) Converting player experience to content, [UTUTU Co, Ltd. and MYSQ 2005]
(5) Making experience transferable by converting to media form

1 and 2 are characteristics of the creative process. Simple operation is extremely important so that players can enjoy using MYSQ, even if for one time only. While maintaining ease of use by enabling multilayered and expressive interactions, MYSQ can provide more replay value. 3 is the social interaction that occurs during the creation process. MYSQ game play is designed on the basis of a collaborative concept that encourages players to help each other to create content that is otherwise not realizable for single players. Items 4 and 5 are social interactions that result from the products of creation. By combining the player's game experience and video footage in one package, MYSQ aims to enrich the contents' trade-ability. This idea arises from the assumption that content that can be identified with concrete facial or other body features enables more communication than content that is not identifiable with specific bodily features. By converting the resulting content to mobile movie form instead of leaving it as a time/space dependent experience, it can be communicated regardless of time and space.

In the next section we review dance-type entertainment systems for multiple players

from the viewpoint of characteristics 1 and 2, and review the entertainment systems that focus on communication arising from content creation by taking characteristics 1 to 5 into account. In Section 3 we discuss MYSQ, its concept, characteristics, and its system. Section 4 is an analysis of player critique. In the evaluation experiment, we use a heuristics based on Csikszentmihalyi's FLOW theory to analyze the system's effectiveness and its entertainment value as a game.

2. RELATED WORK

2.1 Player-Participating Dance Entertainment Systems

There are systems that sense the player's dance movements, and use movement parameters to control video effects, enabling the player to perform together with the resulting video [Meador et al. 2004; Toeplitz 2005]. .These systems are typically customized for professional dancers and lack the ease of use that a naive player would require.

On the other hand, commercial games such as Dance Dance Revolution [Konami Corporation 1998] and Para Para Paradise [Konami Corporation. 2001] do not require any dance experience. In Dance Dance Revolution, the player responds to graphical objects on the screen by performing steps, and in Para Para Paradise, the player waves his arms. Sound is played as a result of these interactions, but while control is extremely simple, interaction is limited to on/off control and the level of expression is not robust. In an attempt to encourage potential players, the player is displayed in a public space via the design of the booth as an L shaped semi-closed space. .But this causes a problem, since the player may be made nervous and tense by an unwelcome audience.

In a therapeutic interactive painting system, called Body-bursh [Ip et al. 2003], the player's position is mapped into 3D space via 3D blob tracking, converting the parameter into sound and graphics. In the painting mode, a line is displayed on the wall showing the player's motion trajectory, with colors and shapes changing depending on the player's acceleration. By using the player's motion as a 3D paint brush, this system accomplishes a simple operation interface, and via motion data and velocity/acceleration, the interaction becomes multilayered and expressive. However, the experience is single player only, so there is no multiplayer collaboration or communication.

MYSQ aims to reduce a cumbersome physical interface, and focuses on extracting the parameter with a camera only. MYSQ aims to be a multilayered interactive system by using the parameters it obtains; it also aims to be a system where, regardless of skill level, a naive player or an experienced dancer can both enjoy the experience.

2.2 Entertainment Systems Focusing on Content Creation and Interplayer Communication

There are many studies of systems that focus on content creation and interplayer communication. Moo-pong [Wada et al. 2005], an original kaleidoscope generation system, uses a proprietary device implanted in a ball with RFID tags that records a video of its environment. Several balls of this type are thrown into a viewing box, enabling the player to view the combined video via kaleidoscope. By alternating the thrower, or the person who manipulates the cameras, and the viewer, who views the kaleidoscope, players collaborate and communicate through the process of content creation. However, the video is recorded on the client's PC, and video playback requires proprietary hardware, so this system is not generic and not in an easily trade-able form. However, because the video is put into a kaleidoscope it has little trade value and is not valuable enough to keep (even though the video enables communication). There is also the problem of the material losing its characteristics due to the process of mixing.

Moving Pictures [Vaucelle et al. 2005] is a video creation and editing system targeted towards promoting collaboration among children. Moving Pictures attaches tokens with ID tags and records video to a proprietary device with built-in RFID reader and camera. This data along with its ID is sent to the computer via wi-fi. The token is placed on a custom-built table with a built-in RFID reader., and there is a display in front of the table where recorded movies can be shown. By changing the order of the tokens, recorded movies and sounds can be edited in an intuitive way. In contrast to moo-pong, which is limited to recording favorite videos, multiple users can collaborate and create scenarios together while editing, thus boosting Moving Picture’s entertainment value considerably. However, there are no effects or other expressive functionalities other than editing, and there is no mentioning of easy methods for trading the resulting videos.

Context Photography [Rost et al. 2005] is a good example of a cellular phone application. In Context Photography, the system acquires the sound and video of its environment via the cellular phone's camera and microphone. The acquired data is used to apply effects to images from the phone's camera. This application has been customized for cellular phones; it is easy to create, save, and trade the resulting graphics.

Print Club [ATLUS Corp. 2005] is an example of a commercial game. In Print Club, a player takes a photo(s) with friends using a camera built into a booth. After taking the pictures, the player uses painting and other editing functions, applies graphic processing, and prints out the results as stickers. Print Club is easy to use and the painting and editing functions allow robust expression. Two players can play at the same time, and thus enjoy collaboration while making the stickers. Furthermore, the images of the players are also material for the product (the stickers), which can be given to people as a fun way of getting to know them, and equally to receive them from other people. The

value of such an item as a communication tool is very high. And due to its sticker form, Print Club products have very high trade-ability. In recent versions [Omron Entertainment 2005], players have the option of sending the resulting images to cellular phones directly, making distant communication possible.

MYSQ inherits many of the characteristics of Print Club, but instead of images uses a video of the process. The strength of video as a media is the time axis. By valuing the process of creation and packaging the players' experience, MYSQ aims to create highly trade-able content. To make saving and trading easier, MYSQ uses mobile devices (mainly cellular phones).

3. MYSQ

3.1 Overview

MYSQ uses a 2.5 cubic meter booth, applying video and audio effects from data acquired by sensing the player's motions (i.e., by detecting the movements of the player's foot and arm). The movements of the foot are tracked by sensors l at the bottom of the booth; arm motions are detected via a custom-made MYSQ ring and image analyzer. The parameters for foot movements are used to select the video and audio effects to apply; the parameters for arm movements are then used to determine how video and audio effects are to be applied. The source data for image analysis (to sense arm movements) is movie data captured from a video camera located inside the booth. The player is able to create an original music video clip by using these facilities. The completed movie is uploaded to the server, and a QR code [Denso Wave. 2005.] representing the uploaded clip's URL is displayed on the screen. The player uses a handheld device or cellular phone to download the uploaded video clip by using the QR code. Once the QR has been captured and converted into URL format, the player can send and share the movie clip with friends. Figure 1 illustrates the flow of events from play start to finish.

The booth (Figure 2) is 2.5 cubic meters, with a 2.2 m-wide by 1.8m-long by 2.5m-high player space inside. There is a moving panel inside the booth that holds both the display and the camera. It also holds the start button and an up/down button that the player uses to configure the camera and display position to suit his or her height. The

Fig. 1. Game flow of MYSQ, from start to end.

Fig. 2. Images of MYSQ (left: outside, center, right: inside).

up-button activates the motor to move the panel up, and vice versa. In front of the player space there are lights that flatter the player's skin tones. The lights are 106,400 kelvin florescent lamps (the same lamps used in photo studios). On the booth's outside wall are 18 LED lamps placed behind panels. The LEDs are grouped into 6 groups with 3 lamps each, and each of the 6 groups is linked to the 6 -foot on/off switches.

3.2 Concept

The concept of MYSQ is the realization of communication through the process and product of content creation. This communication can be divided into three categories.

The first is the communication between players inside the booth. MYSQ game play is not limited to one player, but is open to collaborative game play by multiple players. For example, one player can choose what video effect to apply, and another may determine the parameter for the effect. Another example is to simply have one player determine and control one-half of the effect, while another one controls the other half; but these are merely some possibilities in the design phase. It is important to leave plenty of leeway for players to create their own styles of collaboration. Collaborative elements in system design are discussed in detail in Section 3.3.3.

The second is the indirect communication between players inside the booth and viewers outside the booth. Ideally, a booth should be designed as a closed space, so that the players can concentrate on game play. However, there are two ways to show the outside viewers what is happening inside the booth. By showing outside viewers the inside of the booth, we aim to raise their curiosity and their desire to play.

As the video effects are applied and fed back into the monitor inside the booth, the same video is also fed to a monitor outside the booth, to enable the viewers to see what the end result could look like. The motions of the player's foot on the switch are shown to the outside viewers via the LED panels, allowing them to view indirectly the happenings inside the booth by watching the LED panels blink along with the player's actions. By receiving both direct and indirect information, the outside viewer can guess the relationship between the LED (foot switch) and the resulting video, increasing curiosity further.

The third is the nonsimultaneous communication between the player and whoever the player shares the video with. The player's MYSQ experience is saved as video, and can be shared via cellular phones and mobile devices. By sending the URL acquired from the QR code to friends, the movie can be shared regardless of physical distance. By using the player as the source material, the video is not limited to a one -time experience, but may be shared with others; instead of using abstract graphics to recall memories, it is more effective to use the player's face or body.

3.3 Characteristics

MYSQ consists of the following five characteristics:

(1) simple operation (characteristic of creation);
(2) multilayered interaction (characteristic of creation);
(3) collaboration (characteristic of social interaction during the process of creation);
(4) converting player experience to content (characteristic of social interaction from the creative product); and
(5) making experience transferable by converting it to media form (characteristic of social interaction from the creative product).

3.3.1 *Simple Operation*. Simple operation is ease of use in creating the content, that is, is the original music video clip. MYSQ captures all video and audio processing parameters from the movements of the player's body. By sensing those body parts that are used most frequently in daily life (i.e., the arm and foot), even first time players can use MYSQ with ease. Video effects can be applied by simply moving around, enabling players with no video editing or graphics knowledge to create videos without stress. Sensing is discussed in more detail in Section 3.4.1.

3.3.2 *Multilayered Interaction*. MYSQ aims for multilayered interaction that is expressive and robust. Multilayered interaction is realized by the combination of the video effects via foot sensors and the effects parameters assigned to the player's arm motions. Each foot sensor is assigned an independent video effect and sound track. The effects parameter is derived from absolute arm movements via image analysis. By binding motion and the temporal axis, we can make effective use of the temporal characteristics of video. By using many instances of these effects in conjunction, it is possible to create multilayered effects. Parameter extraction and individual effects processing are discussed in Section 3.4.2.

The basic use of effects is simple, and a first time player can create expressive video clips with ease through the use of body movements. If a player has video editing or video production experience, the player may opt to use the music as cue to apply effects in sync with the music, or use other standard video editing techniques. If a player is an

experienced dancer, he or she may use dance as a primary form of expression, and only apply effects to accent the dance.

3.3.3 *Collaboration*. MYSQ is a multiplayer system. There are two features that assist in collaborative game play. One is via design of the booth, whereby a player is enclosed by walls so that he or she can concentrate on game play without being hindered by viewers. By placing the foot sensors very close to each other, players are forced to come into close contact, and as a result have a greater tendency to communicate.

The second feature is through software design. Collaboration in previous entertainment systems was similar to that in the analogue world - a loop where a player's action inspires another player's reaction. MYSQ combines the collaborative hardware features with the player's input and arrives at collaborative computing. In detail, multiple players' arm motions are summed on a per frame basis and the effects are controlled by the sum of all the player's motions. The algorithm is discussed in Section 3.4.1.

3.3.4 *Converting Player Experience to Content.* The video clip created by MYSQ is the player's experience packaged into video form, which is essentially the player's original music video clip. The player becomes the source material, and the process through which the player applies effects via movements of his/her body is the core content of the experience. By retaining the player's experience and recognizable visual features, the resulting video clip becomes more valuable and is more likely to be traded. With still images the result is visible, but not the process. So it is hard to imagine the emotions that went into creating the result. In a MYSQ movie, the entertainment value of the process is used in conjunction with the player's recognizable visual features, making it possible for the viewer to recall the emotions that went into creating the video clip.

3.3.5 *Making Experience Transferable*. The MYSQ experience can be used as a new communication tool by converting it into media form, thus making it easy to trade and share. Here, conversion from experience to media means conversion of the player's experience in creating the movie (process) to a saved and trade-able form. Trading is mainly via cellular phones. It is possible to create videos on cellular phones, but that requires literacy, and the cellular phone's video editing functions are very limited. In contrast, MYSQ does not require any literacy, the video can be created in an entertaining context, and traded and saved using cellular phones (an everyday item). Saving and trading are discussed in detail in Section 3.4.3.

3.4 System

The MYSQ System (Figure 3) can be divided into the content-creation process and the communication process that occurs after creation. The content-creation process is divided into sensing and feedback. In sensing, input from the foot switches are processed

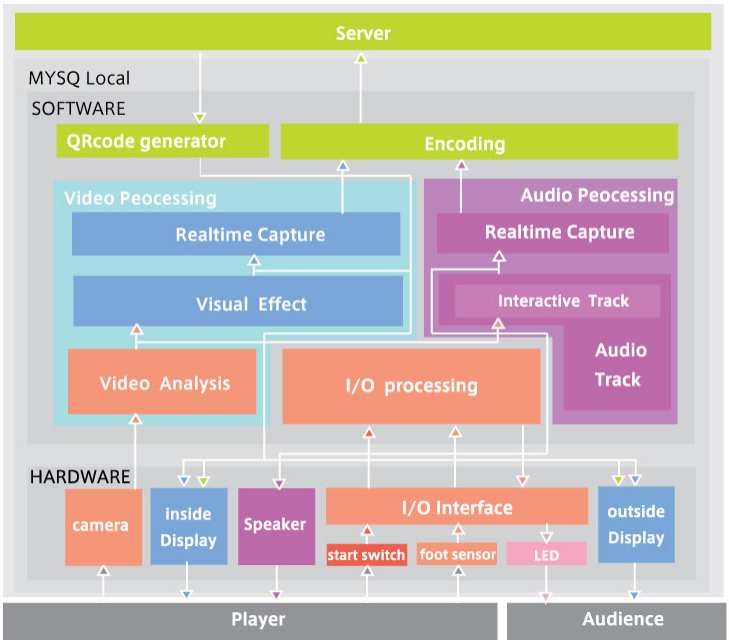


Fig. 3. System architecture.

using an I/O processing module. Camera input is processed through a video analysis module. In feedback, video and audio processing is applied using the parameters acquired through the sensing phase, and the result is fed back to the monitor inside the booth, LED, and speakers. During the communication process, the system performs server communication and displays the QR code for the file URL on the monitor inside the booth. The following sections l discuss each of the processes in detail while showing the implementation of the characteristics mentioned in Section 3.3.

3.4.1 *Sensing*. MYSQ senses the player's foot and arm motions. In order to make the interface easy to use, MYSQ opts for an interface that does not restrict the player's motions. To make MYSQ a multiplayer system, it uses a collaborative computation feature.

For foot-sensing, we assume a dance step-like motion. To sense the player's steps, there are 6 foot switches (Figure 4) on the booth's floor, which has a 2kg on/off threshold. The foot switch turns on when it is stepped on. The on/off information is sent through an I/O interface and is converted to a 1- bit 1/0 signal via the I/O processing module.

To sense the player's arm motions, MYSQ uses a custom-made MYSQ ring (Figure 5) and a digital video camera. The MYSQ ring consists of a high intensity red LED light, a power switch, and batteries. It is assumed that the users will wear the MYSQ ring on

Fig. 4. The 6 foot sensors.

their palms. The weight is about 100g, and does not encumber the player at all. By applying image analysis to the video stream acquired from the digital video camera located at the front of the booth, the application detects the MYSQ ring's LED light and color information and plots them on a 2-plane.

A display is installed inside the booth so that the player can view the resulting video and interact intuitively. By feeding back the resulting video in real time, the player can kinesthetically learn how the motion affects the video. By default, the video is mirrored and fed back to the display.

To extract the LED location information on the MYSQ ring, the LED's color (C) and brightness (L) are extracted. The LED color is extracted in RGB format (R,B,G) from a still frame captured from the digital camera and stored in a buffer. Brightness information is obtained by converting the still frame used for color extraction into 255 color grey-scale and extracting the ring's brightness as an alpha value (α). The brightness alpha value is also stored in a buffer. Using these two values and finding the exact location of the ring, motion is detected by calculating the difference between consecutive frames.

The center of the monitor is the origin (0, 0) and the plane of motion is written as an XY axis (-100<X<100, -100<Y<100). For any specific frame T_n, the location of the ring is written as $P_n(X_n,Y_n)$. The location of the point that matches the color information stored in buffer $C(R,B,G)$ is written as $C_n(X_n,Y_n)$ on plane XY. The location of the point that matches the brightness information stored in buffer $L(\alpha)$ is written as

Fig. 5. The MYSQ ring (left: off; right: on).

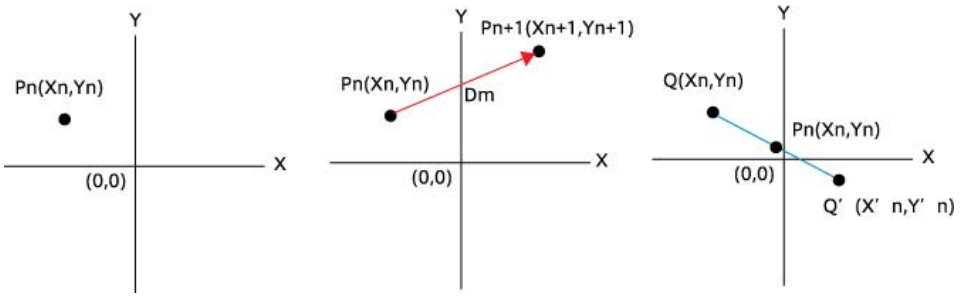


Fig. 6. Example of ring position and motion extraction: (left (a): extracted point Pn; center (b): motion DM; right (c): multiple points).

$L_n(X_n, Y_n)$ on plane XY. When P_n fulfills expression (1), P_n is the location of the ring (Figure 6a).

$$P_n(X_n, Y_n) = C_n(X_n, Y_n) = L_n(X_n, Y_n) \quad (1)$$

$T_{n+1}(X_{n+1}, Y_{n+1})$ is the frame M(msec) after frame T_n. Motion D_M occurring during the M(msec) can be written as in Eq. (2). The ring position and motion derived from Eq. (2) using Euclid distance (Figure 6b). The derived position and motion are used in conjunction with on/off information from the foot switches to control the video and audio processes.

$$D_M = \sqrt{(X_{n+1} - X_n)^2 + (Y_{n+1} - Y_n)^2} \quad (2)$$

When more than one point is detected during M(msec) such as during collaborative game play, the average of the points is used (Figure 6c). Motion is then computed by the difference of the average points between frames. When two players are playing, the motion and position of both players can be used for video and audio processing. If both players move simultaneously, they can control the effects in harmony. If the players are not synchronized or are not in harmony, the effects cannot be fully controlled. However, it may be quite entertaining to enjoy the unpredictability of dissonant motion.

3.4.2 *Feedback*. The system uses the foot switch on/off information, ring position, and motion acquired through the sensing process and uses the information for video and audio processing. The foot switch is further used to control the LED panels located on the outer wall of the booth. For multilayered interaction, a robust interface is used to enable complex expressions, depending on the player, while keeping the basic interactions simple. The player's experience becomes the core content through the use of the player's recognizable features, which accurately reflect the player's experience inside the booth.

MYSQ applies real-time video effects to the user's video captured from the digital video camera located inside the booth. The video effects are assigned to the six foot

switches on the floor of the booth. The effects' parameters are determined by the motion of the ring (D_M), which is the difference between the position of point $P_n(X_n, Y_n)$ and $P_{n+1}(X_{n+1}, Y_{n+1})$ from two consecutive frames T_n and T_{n+1}, M(msec) apart.

MYSQ is configured to use a M value of 200msec. This value was derived from the threshold of latency to multimodal stimulation [Sweetser and Wyeth 2005]. D_M is calculated from frames T_n and T_{n+1} using 200msec intervals, and D_M is applied during the 200msec interval between T_{n+1} and T_{n+2}. The resulting movie frame rate averages 15fps on a PowerMac G5 Dual CPU 2.5GHz, Radeon 9800 Pro 256MB, Memory 2GB.

In Table I, we list the six difference video effects offered in MYSQ, (i.e., name, effect description, relationship to user's motion, and sample graphic).

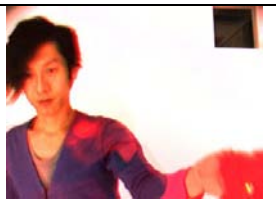
These effects are not only controlled by an on/off toggle switch, but also by the position and motion of the MYSQ ring. By using different kinds of motion, the player can create many different kinds of interactions. Effects can be used simultaneously; the goal is for users to find their favorite, original effects by combining them.

The effects are assigned to foot switches by row, according to similarity. The two foot switches in front are assigned to Star and Heart effects, both of which generate shapes. The mid-row is assigned to Pinky Line and Pink Contrast, both are pink effects. The last row is assigned to Ghost and Kaleido, both of which use the player's image as an integral part of the effect as the source image is dynamically updated. In addition to these default effects, there are several others such as particle or strobe effects. The special effects are triggered when a specific sequence of steps has been entered; the special effect is then superimposed on the normal effects.

MYSQ has eight different music sound sets that play during game time. The sound sets are divided into styles such as hip hop, R&B, house, and so on. While the game loads, a sound set is chosen at random. A sound set consists of seven tracks: the first track is a basic rhythm section and bass; the remaining six are assigned to the foot switches. When the game starts, track 1 (basic track) is played at full volume. When the player steps on a foot switch and the I/O processing module receives the digital signal from the foot witch, the assigned track fades in with a 500msec curve. As the player steps off the footswitch, the track fades out with a 1000msec curve. The fade in/fade out are inserted to reduce aural discomfort resulting from sudden changes in volume. The fade in time is set shorter than the fade out time in order to maintain quick feedback response to the user's actions.

The foot switch on/off information is reflected on the LED panels located on the outside wall of the booth. Each foot switch is linked to a specific group of LED panels. The LED panel is turned on when the linked foot switch is turned on. There are 25 * 25 high intensity red LEDs facing in all directions. The red light is made pink by shining it through a semi-transparent white acrylic panel. The viewers surrounding

Table I. List of Video Effects

Name	*Effect description*	*Relationship to user's motion*	*Sample graphic*
Star	Real-time star generation	Stars are generated at the ring position. The more the player moves, the more stars are generated, and vice versa.	
Heart	Real-time heart- image generation	Hearts are generated at the ring position. The more the player moves, the more hearts are generated, and vice versa.	
Pinky Line	Real-time pink line generation	A pink vertical line of varying width is generated and moved horizontally. The more the player moves, the more lines are generated, and vice versa.	
Pinky Contrast	Real-time change in contrast of source image	Using the player's amount of motion, the RBG values are attenuated to make the image appear pink. The greater the motion, the quicker the source image is tweened to a pinkish tint.	
Ghost	Real-time ghost- effect generation	The source image is cloned onto four separate layers. The more the player moves, the faster the four layers move apart from each other in the horizontal and vertical directions.	
Kaleido	Real-time kaleidoscope-effect generation	Using the player's amount of motion, the source image is rotated and scaled. The more the player moves, the faster the rotation is applied and the smaller the image becomes.	

the booth can indirectly observe what is going on inside the booth by watching the LED panels. During standby, the LED panel groups fade on and off, in a random sequence. The standby sequence allows MYSQ to appeal to potential players, showing them that the booth is ready for play.

3.4.3 *Communication.* The resulting movie is uploaded to a server and can be downloaded by mobile devices. For easy trading, cellular phones and QR codes are used for downloading and cellular phone mails are used for sharing. This section describes the process and implementation.

The processed video and audio fed back to the player during play time are captured in real time within the application. After the game ends, the stream is encoded into MPEG 4 3gp/3gp2 (cellular phone video formats) via Quicktime. The Perl script generates a random four-digit number that is used for the filename. The Perl script also generates a html file that holds the encoded video. The script then generates a QR code with the URL to the html file that holds the encoded video. After encoding, the html file and MYSQ movie are uploaded to a server. The generated QR code is displayed on the monitor inside the booth. The player uses a mobile device and captures the QR code and downloads the movie file without having to input the movie URL or mail address by hand. For players that do not have QR code capability, the access number is displayed as well, so the player can input it by hand. The player can then use commonplace cellular phone functions to share the URL obtained through QR code, or send the QR code image. Currently, there is a test cellular phone located beside the monitor for players who have incompatible cellular phones and for testing.

In the current version of MYSQ, the movie is limited to the least-average file size of 250KB that can be downloaded by cellular phone. Based on this, the play time is set to 30 seconds. Capture occurs in real-time, encoding takes 15 seconds, and uploading takes roughly 5 seconds. The time it takes from play start to the QR code (URL) being displayed is about 1 minute. To shorten the user's wait time, all video is encoded with the same setting regardless of cell phone capability.

4. EVALUATION

4.1 User Study

A user survey was conducted to evaluate the MYSQ system as a game. The survey items were created using Penelope's Game Flow [Sweetser and Wyeth 2005] which is an expansion of Csikszentmihalyi's [1990] FLOW theory, a model consisting of eight items that measure a subject's feelings of enjoyment. It was created by a psychologist called Csikszentmihalyi, who interviewed several thousand subjects during a time span of more than ten years. Penelope et al. [Sweetser and Wyeth 2005], keeping conventional game theories in mind, based their research on the FLOW theory and realized a system to evaluate the entertainment value of a game. Though much needed, such an evaluation

system never existed before. Penelope's Game Flow is broken down into the following eight features.

- *Concentration*: Games should require concentration and the player should be able to concentrate on the game.
- *Challenge:* Games should be sufficiently challenging and match the player's skill level.
- *Player skills:* Games must support the development and mastery of the player's skill.
- *Control:* Players should feel a sense of control over their actions in the game.
- *Clear goals:* Games should provide the player with clear goals at appropriate times.
- *Feedback:* Players must receive appropriate feedback at appropriate times.
- *Immersion:* Players should experience deep but effortless involvement in the game.
- *Social interaction:* Games should support and create opportunities for social interaction.

The game flow mentioned above has been proven applicable to online games, and cannot be applied accurately to a system like MYSQ where game play is based on creation. As a result, original questions were created based on the Game Flow features above. Table II shows the questions, their aim, and the corresponding game flow feature.

The survey was conducted on first-time users of MYSQ, right after game play in the first two weeks of December 2005 at KDDI Designing Studio [2005] where MYQ was set up. There were 110 subjects that had valid responses during the survey period. The age distribution was 2.1% below 10 years old, 40.0% 10-19 years old, 46.3% 20-29 years old, 9.4% 30-31 years old, and 2.1% 40-59 years old. The average age of the subjects was 18.72 years old; 35.5% were male and 64.5% female. For this evaluation, the results of the survey were not sorted, but all items were evaluated to obtain an all inclusive feedback of the system. The results are illustrated in Figure 7.

For question (i), "Could you concentrate on game play?" 38% answered "could concentrate very well;" 51% answered "could concentrate quite well." All in all, 89% answered positively.

For question (ii), "Do you have video editing or dance experience?" 10% answered "have had video editing experience;" 25% answered "have had dance experience;" 12% answered "have had experience in both fields." Of all those who answered that they had some experience, 4% answered "could make use of video editing experience;" 25% answered "could make use of dance experience;" and 8% answered "could make use of both experiences." To question (iii), "Which experience could you make use of?" For

subjects with only video editing experience, 18% answered "could make use of video editing experience;" while 82% answered "couldn't make use of video editing experience." For subjects with only dance experience, 33% answered "could make use of dance experience;" while 67% answered "could not make use of dance experience." Of subjects with experience in both fields, 31% answered "could make use of experience in both fields; "39% answered "could make use of dance experience;" 30%

Table II. Graph 1. User Survey Questions, Aim, and Corresponding GameFlow Feature

Question	*Aim*	*Game Flow*
(i) Could you concentrate on the game?	Could the subject concentrate during the 30 seconds of game time without becoming bored.	Concentration
(ii) Do you have video editing or dance experience?	Does the subject have video editing or dance experience, and if yes, could they effectively use it during game play. If the subject did not have any experience, did they enjoy the game regardless of experience. The question aims to check if the system could cater to both experienced and non- experienced players.	Challenge, Players Skills
(iii) If you answered a (had experience in both fields), b(have had video editing experience) and c(have had dance experience) for question ii, could you apply your experience to MYSQ?		
(iv) If you answered d (no experience in either field) in question (ii), did you enjoy the game play?		
(v) Did you feel you were controlling the video and audio?	Check if the subject could freely control the effects.	Control
(vi) Did you understand the concept of making a cool/cute music video clip to send to your friends?	Check if the concept was easy to understand.	Clear Goals
(vii) Did the visual effects appear in sync with your body movements?	Check if the visual effects were applied correctly	Feedback
(viii) Did you feel immersed during game play?	Check for immersion	Immersion
(ix) Did you try multiplayer?	Check if emphasis was on computational collaboration or analogue collaboration.	Social Interaction
(x) If you answered yes to question (ix), did you concentrate more on dance or on controlling the effects?		
(xi) Did you download your MYSQ movie?	Check if the subject downloaded the movie and shared it with friends. Question (xiii) was made on the assumption that (xi) and (xii) show correlation.	
(xii) If you answered that you downloaded using your own cellular phone or downloaded using the test cellular phone to question (xi), did you send your MYSQ movie to your friends?		
(xiii) Have you ever sent movies to your friends prior to experiencing MYSQ?		
(xiv) Is a MYSQ movie valuable to you because, like Print club, it is an image of yourself?	Check the effectiveness of using the subject's experience as content and the effectiveness of content as a communication tool.	
(xv) Is a MYSQ movie useful as a communication tool?		

answered "could not make use of experience in both fields." From these survey results it is apparent that subjects used dance experience more than video editing experience. For question (iv), "Could you enjoy the game play?" directed towards subjects with no experience in either field, 46% answered "enjoyed very well;" 49% answered "enjoyed quite well;" 95% of all answers showed a positive response. For question (v), "Did you feel like you were controlling the video and audio?" 12% answered "felt very well in control;" 50% answered "felt quite well in control;" 62% of all responses showed a positive response. For question (vi), "Could you understand the concept of making a cool/cute music video clip to send to your friends?" 12% answered "could understand very well, 51% answered "could understand quite well", of all responses, 63% showed positive response. For question (vii) "Did the visual effects appear in sync with your body movements?" 19% answered "felt very much in sync;" 53% answered "felt quite well in sync;" and 72% of all responses showed a positive response.

For question (viii), "Did you feel immersed during game play?" 32% answered "felt very immersed;" 44% answered "felt quite immersed;" 76% of all responses showed a positive response.

For question (ix), "Did you try multiplayer?" 86% answered "yes;" of those that tried multiplayer, 53% answered "concentrated on both controlling the effects parameters and dancing;" 33% answered "concentrated on controlling the effects parameters;" 86% of all responses showed support for computational collaboration.

For question (xi) "Did you download your MYSQ movie?" 29% used their own mobile devices to download their movies and 37% used the test cellular phone that is a part of the booth. Together, 66% of the subjects downloaded and viewed their own movies. For question (xii) "Did you send your MYSQ movie to your friends?" 2% answered "already sent movie;" 52% answered "planning to send the movie."

Question (xiii) was used to check for a correlation between questions (xi) and (xii), as to whether subjects had ever shared cellphone movies. Of the 35% subjects that had shared cellphone movies prior to game play, the correlations to question (xi) were that 42% downloaded their MYSQ movie with their own cellphones and 37% downloaded the movie with the test cellular phone that is part of the booth. In comparison to all the other responses, there is a 13-point increase in "downloaded movie with their own cellphone" and a 13-point decrease in "did not download the movie." For question (xii), 58% answered "planning to send movie," showing a 6-point increase. From these results, it is apparent that subjects with prior cellphone movie-sharing experience are more likely to actively use their cellular phones to download and share their movies and show a positive correlation. Of the 65% that had no experience of sharing cellphone movies, the correlations to question (xi) were that 23% downloaded their MYSQ movie with their own cellphones and 37% downloaded the movie with the test cellular phone that is a part of the booth. In comparison to all the other responses, there is a 6-point

Fig. 7. The results of questions (i) to (xv).

decrease in "downloaded movie with their own cellphone" and a 6-point increase in "did not download the movie." For question (xii), 46% answered "planning to send movie,"showing a 6-point decrease. It can be concluded that subjects with no experience in sharing cellphone movies have a stronger tendency to not download their movies and are not very active in sharing them, showing a negative correlation.

For question (xiv), "Is a MYSQ movie valuable to you because it is an image of yourself?" 34% answered "yes, very much so;" 49% answered "yes, quite so;" 83% of all responses were positive. For question (xv), "Is a MYSQ movie useful as a communication tool ?" 38% answered "yes, very much so;" 55% answered "yes, quite so;" 93% of all responses were positive.

For free written responses, other than a few comments requesting improvements in graphics resolution, most gave a positive response.

4.2 Discussion and Future Work

Of the 16 features in the survey, Concentration, Control, Clear Goals, Feedback, and Immersion received positive responses in a majority of the responses.

For the Challenge and Player Skills features, results show that the majority of the subjects with no video editing, dance, or any experience related to controlling MYSQ, showed a positive response. For subjects with either video editing or dance experience, results show that dance experience was more relevant to MYSQ game play. However, the results do not show adequate positive response from experienced subjects, since MYSQ does not have functions that directly relate to experience in dance or video editing. Thus users with such experience could not make use of prior experience in a concrete way. In terms of video editing experience, the subject's previous experience could be enhanced by incorporating more standard video effects such as cuts and fades. In terms of dance experience, the subject's previous experience could be enhanced by incorporating standard dance moves and /steps into game play.

Results show that for questions concerning the Social Interaction feature, the majority of the subjects showed a positive response to collaboration. However, only a little over 50% of the subjects planned to, or actually sent, their MYSQ movies to friends, perhaps because the survey was conducted right after game play. But another possible problem may be the quality of the movie, as several subjects mentioned graphic resolution, the current movie file size may be too small for quality that is good enough for sharing/trading with a wide audience. In the future, movie encoding will be split into encoding for high-end cellular phones and low-end cellular phones, so that movie quality matches that of the cellular phone. From these results, it can be concluded that although there is still room for improvement, MYSQ is entertaining as a game.

The survey questions can also be used to check for the effective implementation of the five characteristics of MYSQ. The *simple operation* feature can be checked with questions (ii) and (iv). From the 95% of respondents who showed a positive response, it can be said that MYSQ is an easy to use system that anyone can use. We think it is important that natural movement was used instead of methods that are cumbersome or unintuitive.

The *multilayered interaction* feature can be checked with questions (v) and (vii). MYSQ is designed for multilayered interaction and expressive content creation. The positive responses to questions (v) and (vii) are 62% and 72%. Although the response shows that implementation was effective, there is room for improvement. Further development will result from the following three views. First, there is the problem of the frame rate. Because the current system applies real-time effects while

capturing in real-time, the final video is only 15fps. This can be improved with hardware and also by programming in a more natural language. Second, there is the possibility of sensing more objects. The current system senses arm movements and foot steps. By expanding sensing to other parts of the body essential to dancing, such as the head or the hips, user expression can be enhanced. However, this will result in an increase in sensing parameters and may complicate interaction, thus creating a trade-off between it and ease of use. Finally, the expressiveness of non-effects content can be improved, which may include superimposing background and foreground items, sound visualization from microphone input, and interaction between visualized objects and the user.

The *collaboration* feature can be checked with questions (ix) and (x). There was a positive response of 86% for computational collaboration, including analogue collaboration; the total positive response for collaboration was 96%. It can be said that MYSQ is successfully causing collaboration. As a method of computational collaboration, the current system uses positions from multiple ring locations and also the average motion across several users. In the future, we would like to develop methods for collaboration that are based on geometric meaning, keeping the mental model in mind—f or example, an effect that activates depending on the shape of the combined silhouettes of multiple users.

The *converting player experience to content* feature can be checked using question (xiv). The system was designed on the assumption that the content's value and trade-ability can be enhanced by using the player's recognizable features and experience inside the booth. The results show 83% positive response and prove the assumption correct.

The *making experience transferable by converting to media form* feature can be checked with questions (xi) through (xiii,) and question (xv). MYSQ is designed to be a communication tool by converting the subject's experience into movie form. For question (xi), 66% of the respondents saved their movie files either on their own cellular phones or the test cellular phone that is a part of the booth. And 54% of them were planning to send their movies to friends. These results are most likely due to graphics quality and robustness of expression. However, for question (xv), 93% of the respondents had a positive response, so it can be said that the communication model and the approach to content design are effective. Trading MYSQ movies relies on cellular phone mail and other functions that people use on a daily basis. In the future, we would like to offer trading/sharing web sites and community sites to further encourage communication.

It can be concluded from the above results that the MYSQ's five characteristics were implemented effectively. And, although there is some room for improvement in graphics quality, expression, and methods of trading, based on the communication of the process and the results of content creation, it can be said that MYSQ is an effective entertainment system

In closing, let us discuss the MYSQ system from a commercial point of view. In the movement of movies toward mobile applications, iPod video and video podcasting had an immense impact. The former encourages carrying movies around and the latter brings back promises similar to those heard in the early days of the internet, “everyone will be able to send out information.” The problem is, obviously, that to do so requires prior knowledge or training; hence, many people are waiting for a system like the blog, which allows them to easily use and send out video. Considering this general direction, it is only natural to conclude that MYSQ, which allows anyone to create movies for mobile application, is a system that fulfills such needs. However, in order to successfully develop a market based on this new culture (as Printclub did), it is important to develop business models based on the characteristics of the movies.

Printclub uses stickers for trading, MYSQ uses cellular phones. Printclub provides enjoyment to users without trading, such as collecting stickers in a notebook and showing it to friends. Due to the digital nature of mobile movies, it is very hard to enjoy a video in the same way. To solve this problem we are considering implementing functionality that allows printing cards of user-selected still frames and QR codes.

In terms of story, Printclub is based on the concept “I want to look my best/cutest for this memorable moment.” On the other hand, while MYSQ;s goal is creating an original MTV movie clip, it falls short of the simplicity of the Printclub concept. To resolve this, we are considering new functionality to allow users to take part in a story, for example expressing feelings for that special someone through a scene in a movie, or experience a virtual space such as dancing on Ayers Rock.

5. CONCLUSION

This article outlined MYSQ, an entertainment system that allows players to create original music video clips (ex: MTV) and share them via cellular phones.

The system's entertainment value as a game was analyzed via user evaluations based on Csikszentmihalyi's flow theory. The survey results were also used to evaluate the five characteristics of MYSQ. Aspects of the system that should be improved include the following: implementing better feedback for players with video editing or dance experience, improving overall video quality, including image quality, and the creation of mechanisms to help communication. By enhancing the MYSQ system, we aim to create a culture of communication through movies on mobile devices.

ACKNOWLEDGMENT

This work was conducted with the help of KDDI Corporation [2003]. We are very grateful to the staff of KDDI Designing Studio, where MYSQ is displayed and used for the evaluation experiment. Special thanks to Kenji Kohiyama of Keio University, who helped with the writing of this article. We would like to express our deepest gratitude for everyone involved in this research.

REFERENCES

ATLUS CORP. 2005. PrintClub (Puri-Kura). http://www.atlus.co.jp/am/printclub.html

BENFORD, S., MAGERKURTH, C., AND LJUNGSTRAND, P. 2005. Bridging the physical and digital in pervasive gaming. *Communi.ACM 48,* 3, 54 -57.

CSIKSZENTMIHALYI, M. 1990. *Flow: The Psychology of Optimal Experience.* Harper Perennial, New York.

DENSO WAVE. 2005. Incorporated QR code. http://www.qrcode.com/

IP, H., YOUNG, H., AND Tang, A. 2003. Body brush: An interface where aesthetic and technology meet. *SIGGRAPH 2003 Emerging Technologies*.

ISHII, H. AND ULLMER, B. 1997. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Ssystems.* ACM, New York.

IWAI, T. AND NISHIBORI, U. 2005. TENORI-ON. *SIGGRAPH 2005 Emerging Technologies*.

IWAI, D. AND SATO, K 2005. Heat sensation in image creation with thermal vision. In *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE2005)*, ACM, New York, 213-216.

KDDI DESIGNING STUDIO. 2005. http://www.kds.kddi.com/topics/index.html.

KDDI CORP. 2003. http://www.kddi.com/english/index.html.

KONAMI CORP. 1998. Dance dance revolution .http://www.konami.com.hk/kma_ac/ac_ppp/ppp.html.

KONAMI CORP. 2001. ParaParaParadise. http://www.konami.com.hk/kma_ac/ac_ppp/ppp.html.

LEVITIN, D. J., MACLEAN, K., CHU, L. Y., AND JENSEN, E. R. 2000. The perception of cross-modal simultaneity. *Int. J. Comput. Anticipatory Syst.* 323-329.

MAERKURTH, C., CHEOK, A., MANDRYK, R., AND NILSEN, T. .2005. Pervasive gaming: Pervasive games: Bringing computer entertainment back to the real world. *ACM Comput.Entertainment 3, 3.*

MEADOR, W. S., ROGERS, T. J., AND O'NEAL, K. 2004. Mixing dance realities: Collaborative development of live-motion capture in a performing arts environment. *Comput.Entertainment 2,* 2.

OMRON ENTERTAINMENT. 2005. SutaparaShot. http://www.furyu.jp/columbus/syohin/new_syohin/main_frame.html

ROST, M., GAYE, L., HAKANSSON, M., LJUNGBLAD, S., and HOLMQUIST, L. E. 2005. Context photography on camera phone. In *Proceedings of Ubicomp 2005*.

RYOKAI, K., MARTI, S., AND ISHII, H. 2004. I/O Brush: Drawing with everyday objects as ink. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, New York, 303-310.

SWEETSER, P. AND WYETH, P. 2005. GameFlow: A model for evaluating player enjoyment in games, *Comput. Entertainment 3,*.3.

TOEPLITZ, K. T. 2005. Caputure. Prix Ars Electronica 2005; Honorary Mention Interactive Art 2005.

UTUTU LTD. 2005. MYSQ (my style so cute !). http://www.ututu.com/mysq/

VAUCELLE, C., AFRICANO, D., DAVENPORT, G., WIBERG, M., AND FJELLSTROM, O. 2005. Moving pictures: Looking out/looking in. In *Proceedings of SIGGRAPH'05 Conference on Abstracts and Applications* (Los Angeles, Calif., July 31 – Aug. 4). ACM, New York.

WADA, Y., USUI, S., URYU, D., AND OKUDE, N. 2005. moo-pong. *SIGGRAPH 2005 Emerging Technologies*.

Received January 2006; revised April 2006; accepted April 2006