U-Drumwave: An Interactive Performance System for Drumming

Yin-Tzu Lin[1], Shuen-Huei Guan[1], Yuan-Chang Yao[2], Wen-Huang Cheng[3], and Ja-Ling Wu[1]

[1] National Taiwan University, Taipei, Taiwan, R.O.C.
{known,drake,wjl}@cmlab.csie.ntu.edu.tw
[2] Digimax Inc., Taipei, Taiwan, R.O.C.
[3] Academia Sinica, Taipei, Taiwan, R.O.C.
whcheng@citi.sinica.edu.tw

Abstract. In this paper, we share our experience of applying the modern multimedia technologies to the traditional performing art in a drumming performance project, U-Drumwave. By deploying an interactive system on the drumming stage, the audience will see augmented visual objects moving on the stage in accord with the performer's drumming rhythms. The creation and display of the visual objects are integrated with the concept of story intensity curve in order to vary the perceptual degree of tension given to the audience during the performance.

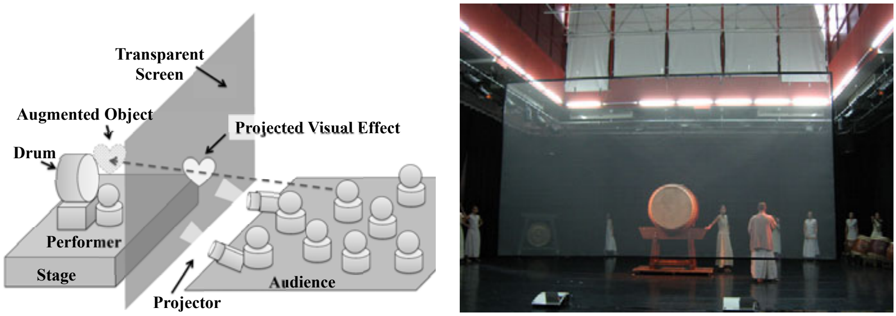
Keywords: Interactive Art, Drumming Performance, Spatial AR.

1 Introduction

U-Drumwave is a project for producing a new style drumming show with the creation of a sense of harmony between the traditional art performance and the modern multimedia technologies. We incorporated the existing interactive and spatial augmented reality technologies into the traditional East Asian drumming performance of U-Theatre[1]. In this show, the U-Theatre artists aim to convey the harmonious relation between humans and the universe. To express the idea, the stage was designed as shown in Figure 1. We set up a transparent screen between the audience and the stage. From the audience's viewpoint, visual contents will be projected on that screen, appearing as augmented objects near to the performer so as to create a perception that there is no boundary between the performer and the world (the projected contents). In addition, the visual effects will be in response to the performer's actions in a synchronized rhythmic manner to express the performer's sense of mutuality and understanding toward the world. To ensure the visual contents' style would conform to U-Theatre's conveyed messages, the contents are jointly created by both the drumming group and our visual design artists through an iterative discussion process. We also incorporate the concept of story intensity curve in the film theory [5] into the

[1] http://www.utheatre.org.tw/

K. Schoeffmann et al. (Eds.): MMM 2012, LNCS 7131, pp. 609–620, 2012.



(a) Overview (b) Actual setup

Fig. 1. The setup of our U-Drumwave stage

design of our visual effects in order to give better comprehension of the idea behind the performer's motions to the audience. The U-Drumwave project was successfully presented to the committees of Taiwan's Council for Cultural Affairs[2] on November 17, 2010, and also granted the government's financial support. In this paper, we will share our experience of running this project and describe our design philosophy and the corresponding implementation details.

1.1 U-Theatre Introduction

U-Theatre is a renowned Taiwanese performance group. It was founded in 1988 by their artist director Ruoyu Liu in a mountain forest in Taipei's Muzha Area. Inspired by Jerzy Grotowski who trains performers in mountains, Liu emphasizes how performers sharpen their bodily sensation and inner awareness. In 1993, after studying meditation in India and Tibet, the drum master ChihChun Huang join the group. Huang requests that before practicing drumming, performers must learn meditation. This training course not only changes the temperament of U-Theatre members but also sets the tone of their performance: athletic drumming and martial arts. Because U-Theatre represents the harmonious encounter between the West and the East in its performances, it is highly evaluated by international artists. U-Theatre has been regularly invited to perform in art festivals of different countries, such as the Netherlands, Italy, Germany, Spain, France, and so forth.

2 Artist's Intentions and Design Strategies

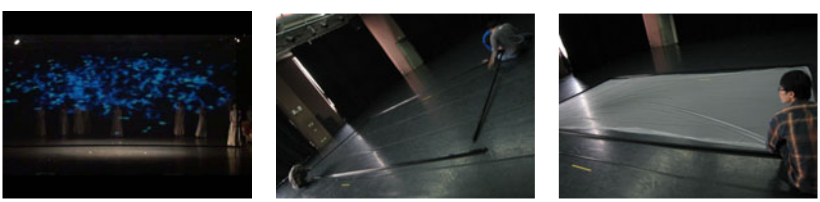
Liu thinks that a show, in essence, is the expression of "an individual's attitude toward life." For U-Theatre members, the attitudes are "living in the moment" and the unification of "Tao[3]" and "Yi[4]". Huang believes that the spiritual states

[2] Taiwan's highest official institution for planning cultural establishments and policies.
[3] A Chinese character generally meaning "way", here means self-improvement.
[4] A Chinese character generally meaning "skill", here means live with art.

during meditating and performing are the same. When the performers drum with meditative minds, each hit on the drums strikes the heart chord of the audience. Contrarily, the loud pounding of drums makes the audience stay calm. Then, the performance becomes a comfort for the audience. In this show, the U-Theatre artists tend to further emphasize the relation between humans and the universe. That is, with open minds, there will be no boundaries among people. Then, there will be no boundary between humans and the world. Humans will develop a sense of mutuality and understanding toward natural beings. To convey the above ideas, our design strategies are listed below.

2.1 Transparent Screen

The transparent screen between the audience and the stage (cf. Figure 1) will help to provide the feelings of no boundary between the performer (humans) and the projected visual contents (the world). From the audience's viewing direction, the contents becomes virtual objects near the performer. There are many means that could create this experience (c.f. Section 6.1). We use a near-transparent projection screen made of gauze. The gauze is coated with special painting materials so that it reflects non-black lights only. Accordingly, black contents projected on the screen become transparent while other-colored contents will be normally shown (c.f. Figure 2(a)). With appropriate lights, stage design and seat arrangement, the audience will not notice the existence of the screen and the immersive experience is created. The advantages of gauze screen are flexibility and portability. For our purpose to "hide" the screen, its size should be as large as possible. Fabricating large gauze screen is easier than making large electronic screen e.g. OLED (Organic light emitting diode). Furthermore, the gauze screen is easy to be decomposed. Its frame and its gauze can be detached (c.f. Figure 2(c)) and the frame can be further divide into short rods (c.f. Figure 2(b)). This makes it easier to be moved among different theatres and more suitable for tour performances.



(a) (b) (c)

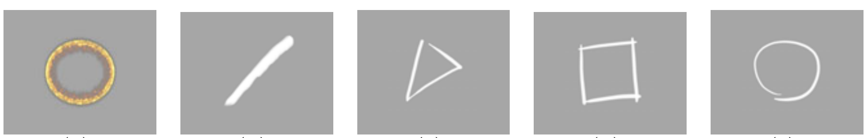
Fig. 2. (a) The gauze screen coated with special painting materials reflecting non-black lights only. (b)(c) Snapshots of our staff's assembling the gauze screen on the stage.

2.2 Interaction

With the assistance of interaction technologies, we can make the visual effect in accord with the rhythms of performers' actions detected from drumming sound. For example, "the time to appear", "the speed of movement", and "the size" of visual effects may correspond to sound events like "timings of hit", "speed of hits", and "strengths of hits", respectively. We regard each visual effect as a virtual character, relative to the real character: the performer. Each virtual character has its own reaction to the performer. Combining all these virtual and real characters forms a story intensity curve (c.f. Section 2.4) that may increase the comprehensibility and attractiveness for the audience.

2.3 Visual Content Design

The visual content are designed through an iterative process. Initially, our visual design artists proposed several designs to the U-Theatre artists. Then, they re-designed the contents according to U-Theatre artists' feedbacks. The above process iterated several times until the both sides have reached a common agreement. As a result, all the visual contents are joint creations of both U-Theatre and our visual design artists. During the design process, we found that the U-Theatre artists prefer plain, abstract, or natural contents, instead of fancy ones. The designed contents are listed below: geometric symbols (Figure 3), Buddhist styled symbols (Figure 4), and simulated natural phenomena video (Figure 5).

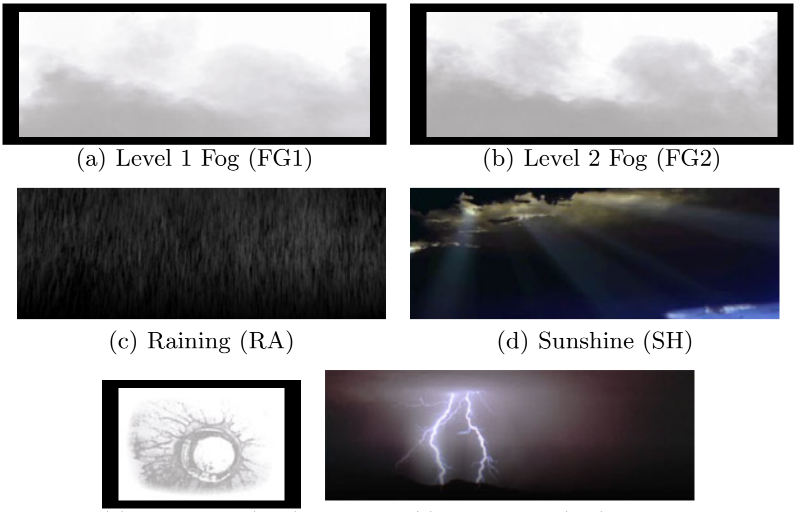


(a) (b) (c) (d) (e)

Fig. 3. Geometric Symbols: (a) Energy Wave(EW), (b) Stroke(ST), (c) Triangle(TR), (d) Rectangle(RE), (e) Circle(CR)

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

Fig. 4. Buddhist Styled Symbols: (a)(b)(c) Part of Abstract Sanskrit Totems (TT), (d) Sutra Video Screenshot (SU)



(a) Level 1 Fog (FG1) (b) Level 2 Fog (FG2)

(c) Raining (RA) (d) Sunshine (SH)

(e) Raindrop (RD) (f) Lightning (LT)

Fig. 5. Screenshots of Simulated Natural Phenomena Videos

2.4 Story Intensity Curve

The term "story intensity" refers to the degree of tension that the audience experienced[5]. A good story that people would enjoy often contains three stages: exposition (EX, beginning), conflict (CO)/climax (CX) and resolution (R)[5]. The story intensity during these stages often follows the shapes like Figure 6(a). In real world, the curve is often jagged (c.f. Figure 6(b)), but the overall shape will still conform to the standard one [5].

To increase the comprehensibility and attractiveness of the performance, both the U-Theatre artist and we carefully arranged the performer's action and the appearing order of the visual contents in order to conform to the standard story intensity curve (c.f. Figure 6(c)). First, the U-Theatre artist adds an exposition section in the beginning of the original pure-drumming to introduce the "interaction" in this show. In the exposition, the performer will act with his motions similar to the projected visual effects, just like he is playing around with the drum. For example, he may use his drum sticks to draw a triangle above the drum with a triangle (cf. Figure 3(c)) showing on the screen at the same time. After that, the performer normally drums loud or soft according to an internal story intensity in his mind. As illustrated in Figure 6(c), the visual contents are also arranged to conform the story intensity curve (including the one hidden in the performer's mind). First, the geometric symbols accompany the exposition section. Then, the Sanskrit totems appear as a bridge between geometric symbols and the sutra book video. As the performer drumming to the climax and suddenly drumming soft, we switch the visual contents to the category of natural phenomena. In this section, we let the worst weather (FG2+RA+LT) match the

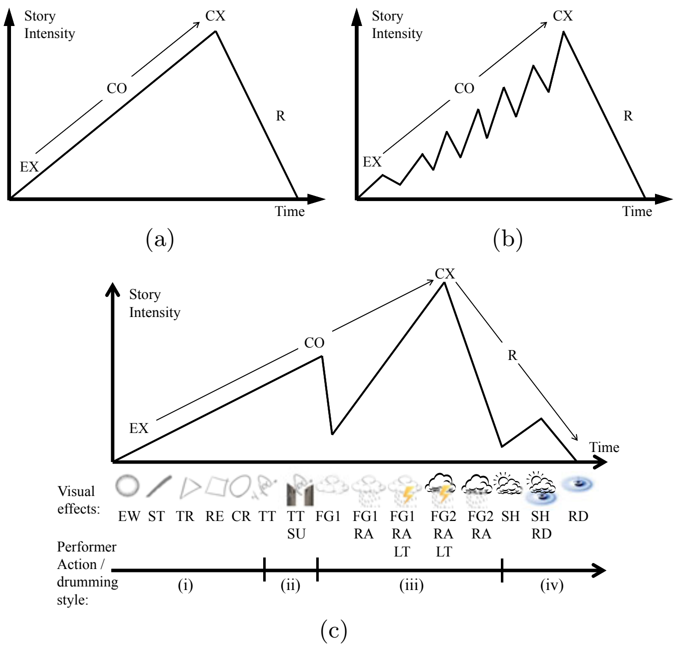


Fig. 6. Standard Story Intensity Curve: (a) Theoretic story intensity curve (b) Real world story intensity curve (c) The Story Intensity Curve of U-Drumwave (i): Testing and playing around with the drum (introducing interactivities) (ii): Drumming to the climax, then suddenly drumming soft (iii): Drumming from soft to loud (climax), and vice versa (iv): Drumming very softly, then stopping drumming, slowly going off stage

climax of drumming. After it clears up and the performer stops drumming, the raindrop effects are shown to decrease the intensity.

3 System Implementations

3.1 Physical Installation

Figure 7 illustrates the physical installation of U-Drumwave. We regard the combination of the blending PC (the red PC) and two projectors (Hitachi CP-A200 3 LCD) as a virtual projector. The main programs running on the control PC (the white PC) analyzes the signals from the audio mixer and then outputs the corresponding visual effects to the virtual projector. During the run-time, the working screen will display the control GUI.

Virtual Projector. Since the gauze screen is wider than the maximum range that one projector can cover, we use two projectors in charge of each half of the projected contents. To reduce the effort of calibration, a blending program

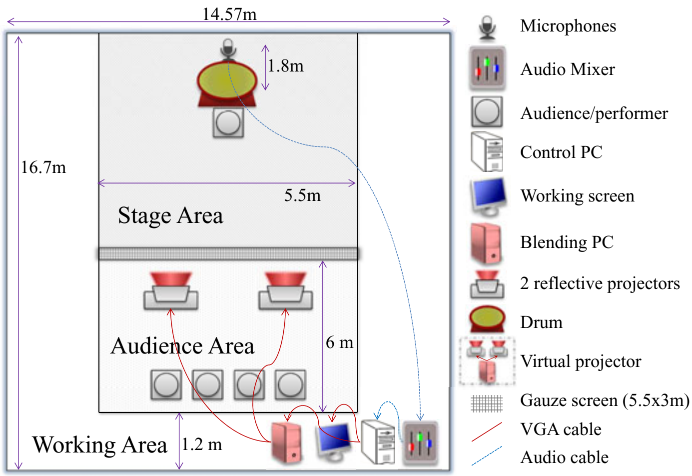
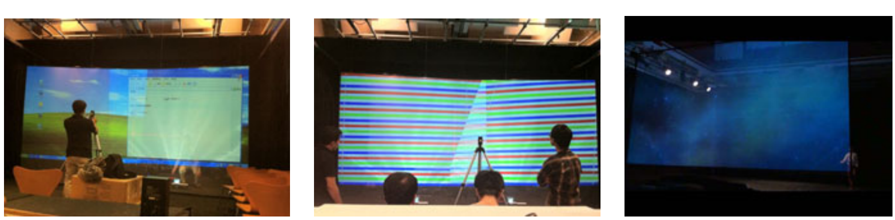


Fig. 7. Semantic Illustration of the Physical Installation

running on the blending PC is embedded as a plug-in of KMPlayer[5], a free video playing software available on the web. There are also tools for helping adjust the overlap ratio of the two halves and tuning the affine transformation parameters of each half (c.f. Figure 8). After adjusting, every video played by the KMPlayer will apply these parameters and show on the gauze screen seamlessly. Since KMPlayer supports real-time playback of frame grabber cards, we let the control PC just output VGA signals as for normal projectors, and KMPlayer on the blending PC plays frames captured from frame grabber card as playing a normal video. For the control PC, the combination of the blending PC and 2 projectors forms a virtual projector.



(a) Before Adjusting (b) During Adjusting (c) After Adjusting

Fig. 8. Adjusting the projection parameters on the stage

[5] `http://www.kmplayer.com/`

3.2 Software Architecture on the Control PC

The software architecture on the control PC is described in Figure 9. We implemented it on the Cycling'74 Max/MSP/Jitter[6] platform. The system can be divided into three parts: audio processing, virtual character, and mixer-like control interface.

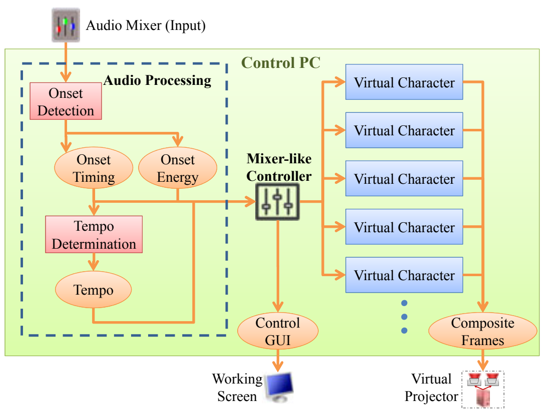


Fig. 9. Control Software Architecture

3.3 Audio Processing

In the audio processing part, we tend to detect the onsets of the input signal. For drumming performance, the onsets can be approximately regarded as beats. Thus, we then determine the tempo information using the timings of the onsets.

Onset Detection. The goal of onset detection is to find the timing and the strength of every hit to the drum. By observation, a hit on the Chinese drum produces a vertical line on the sound spectrogram. As a result, our approach is similar to the "spectra difference" method mentioned in [1] except that we use bark-scaled spectrogram instead and a different peak picking method. The bark-scaled (25 subbands) spectrogram is computed using Tristan Jehan's MSP external "`bark~`" [8]. The onset detection function $f(n)$ is calculated by the sum of the rectified energy difference of each frequency band between successive audio frames. We modified the peak picking method to adapt to this live application. That is, the determination of onsets only depends on previous frames. If $f(n)$ exceeds a certain threshold, we report an onset happened at time n. However, when the performer drums fast, $f(n)$ becomes less discriminative. As a result, two thresholds are used. The total energy in the previous frame $E(n-1)$ determines which threshold to use. As shown in Equation 1, the indicator function $O(n)$ represents the detected

[6] `http://cycling74.com/`

onsets. When $O(n) = 1$, an onset is detected, we output a "bang" message and the corresponding $f(n)$ value reflecting its strength. That is,

$$O(n) = \begin{cases} 1 \, , E(n-1) > T_E^{high} \text{ and } f(n) > T_f^{low} \\ 1 \, , E(n-1) \in (T_E^{low}, T_E^{high}] \text{ and } f(n) > T_f^{high} \\ 0 \, , \text{otherwise.} \end{cases} \tag{1}$$

In addition, the widely used spectral brightness feature[9] is used to reduce false detection caused by other sounds (e.g. human voice). We regard an onset happens if $O(n) = 1$ and $Brightness(n) < 0.08$ because the drum usually sounds bass. The core of the onset detection algorithm is implemented in JavaScript.

Tempo Determination. After obtaining the timing of each onset, we further estimate the tempo (in BPM, beat per minute) of the drumming performance through the reciprocal of onset intervals. An MSP built-in object `sync~` supports the above BPM calculation. However, using `sync~`, the BPM value only updates when it receives "bang" message. As a result, we implement our own patch to immediately update the tempo once the performer starts to slow down. We not only calculate BPM value upon receiving a "bang", but also check the elapsed time from the last onset to now. Once the elapsed time exceeds the last onset intervals, we replace the onset interval as this elapsed time. Subsequently, the output tempo value will gradually decrease when the performer start to slow down until he hits the drum again.

3.4 Virtual Character

The virtual character part deals with the response actions of the visual contents to the audio events. Figure 10(a) shows the structure of one virtual character and its relationships to other components. Users will decide whether a virtual character is allowed to show through the control interface. Once being allowed, the virtual character starts to receive audio processing results. Then, it appears according to the properties of showing based on its settings. For example, the size may be full-screen or varying with the onset energy. The vanish time may be receiving "not allowed to show" or automatic vanishing upon last video frame. The virtual character is implemented as a wrapper object of "`jit.qt.movie`" and "`jit.gl.videoplane`". We choose this combination because it is easier to composite multiple images and videos with varying properties. Besides, the alpha channel works better with "`jit.gl.*`" objects. This makes it easier to show irregular shape objects without revealing annoying rectangle frame and blocking other objects beneath them.

3.5 Control Interface

Figure 10(b) shows a screenshot taken from the control interface of U-Drumwave system. Motivated by audio mixer devices, we design 2 check boxes and one slider

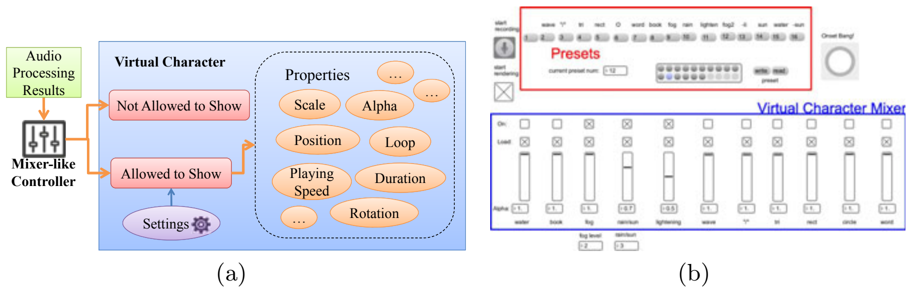


Fig. 10. (a) The structure of Virtual Character (b) Control Interface Screenshot

bar for each virtual character (the bottom part). One check box controls loading the virtual characters into the memory. The other one determine whether the characters are allowable to show. The slider bar controls the transparency of the character's appearance. we also design a presetting part to record the appearance settings, so the users can switch numerous settings of virtual characters simultaneously. To avoid sharp changes between settings, the slider bar will smoothly slides to the target value. "Fade-in/out" effects between the characters are then resulted. The "onset bang" button at the right hand side will light up each time while an onset is detected. If an important event is miss detected, clicking the button to send an "bang" message is a remedy.

4 Live Demo

The U-Theatre artists combined U-Drumwave with their other drum songs to create a 40-minutes show and presented it to the committees of Taiwan's Council for Cultural Affairs on Nov. 17, 2010. Figure 11 presents some snapshots taken during the performance[7]. The government-appointed reviewers highly appreciated the show. They commented like, "Very splendid! Fortunately this is not another tech-show! The technology and the art mix just right and seamless." After the show, the reviewers also enjoy the interactive system. The exposition of U-Drumwave acquired the most interest. Finally, they granted financial support for U-Theatre to run the formal version on August, 2011 in Taiwan.

5 Discussion

The most difficult but also the most interesting development challenge of the U-Drumwave project is the communications between engineers and artists. Our crew comes from various professional domains. We have the drum master, the artistic director, the visual design artists, the gauze screen technicians, the recording technicians and the interactive program engineers. The U-Theatre

[7] Demo video: `http://www.cmlab.csie.ntu.edu.tw/~known/ud/`

(a) (b) (c)

Fig. 11. Snapshots captured from the live U-Drumwave performance

artists often describe things via feelings. These descriptions are vague for the engineers to write programs. Besides, the U-Theatre artist used to impromptus and often change the plot when rehearsing. It is a challenge for the engineers to change the programs as fast according to the artists' requests. The control UI helps a lot to these issues. The artists and the engineers can discussed the flow and the design of the contents with viewing the results on the gauze screen. Besides, it is easily to immediately modify virtual characters' actions, just like a real man changing his gestures. Thus, the resulting performance nicely harmonizes the art and the technology.

6 Related Work

6.1 Spatial Augmented Reality

Spatial Augmented Reality (SAR) is one of AR's sub-domain. SAR does not request users to equip with eye-worn or head-mount display. As Bimber and Raskar[4] mentioned, existing SAR techniques can be categorized into 3 groups according to the display techniques used: "screen-based video see through display", "spatial optical see-through display" and "projection-based spatial dispaly". The first one mixes captured image and virtual object as video and displays on regular monitors. The second one overlays synthetic image on real object through optical techniques, such as optical combiners [3], transparent screen ([10] and our approach) or optical hologram[2]. The last one directly projects image seamlessly on physical object's surface instead of a projection screen[7].

6.2 Interactive Technology for Performance

The development of interactive technology for performance has long been explored. The first representative work was Gordon Pask and Robin McKinnon-Wood's electromechanical system MusiColour in 1953 [6]. Motivated by the concept of synaesthesia, Pask makes color lights adjusted in response to live music. The responded color lights are treated as another performer on the stage. There are many follow up systems. Sparacino et al. [11] gave a widely survey of interactive spaces of performing art. A recent approach was Wei et al.'s Ozone[12], a state-based interactive system for live performance.

7 Conclusions and Future Work

In this paper, we share our experience of harmonizing the technology and the art to produce a show. The artistic intentions are realized through the development of a technology enabled interactive system on the performing stage, including the combinations of a transparent screen, the sound-driven interactive technology, iterative visual content design and the concept of story intensity curve. It is our belief that U-Drumwave has created a new style of performance and set a working example. As for the further improvements, possible directions include enhancing the onset detection accuracy, incorporating timbre recognition, real-time rendered particles, and human action detection with IR or depth cameras.

Acknowledgments. The rest crew of U-Drumwave: Ruoyu Liu, ChihChun Huang, Tai-Huei Lee, MAKA Inc., Chris Wu, Vivi Chen, Eddie Lin, Jade Cheng, Bala Tsai, Chien-Tse Yang, Hao-Ming Chen, and Edward Chang.

References

1. Bello, J.P., Daudet, L., Abdallah, S.A., Duxbury, C., Davies, M., Sandler, M.B.: A Tutorial on Onset Detection in Music Signals. IEEE Trans. Audio, Speech, Lang. Process. 13(5), 1035–1047 (2005)
2. Bimber, O.: Combining Optical Holograms with Interactive Computer Graphics. IEEE Computer 37(1), 85–91 (2004)
3. Bimber, O., Fröhlich, B., Schmalstieg, D., Encarnação, L.M.: The Virtual Showcase. IEEE Comput. Graph. Appl. 21(6), 48–55 (2001)
4. Bimber, O., Raskar, R.: Spatial augmented reality: Merging real and virtual worlds. A K Peters, Ltd. (2005)
5. Block, B.: The Visual Story: Seeing the Structure of Film, TV and New Media. Focal Press (2001)
6. Haque, U.: The Architectural Relevance of Gordon Pask. Architectural Design 77(4), 54–61 (2007)
7. Jacquemin, C., Chan, W., Courgeon, M.: Bateau Ivre: an Artistic Markerless Outdoor Mobile Augmented Reality Installation on a Riverboat. In: ACM Multimedia, pp. 1353–1362. ACM Press (2010)
8. Jehan, T., Schoner, B.: An Audio-Driven Perceptually Meaningful Timbre Synthesizer. In: International Computer Music Conference (2001)
9. Lartillot, O., Toiviainen, P.: MIR in Matlab (II): A Toolbox for Musical Feature Extraction from Audio. In: The International Society for Music Information Retrieval Conference, pp. 237–244 (2007)
10. Laser Magic Productions: Holograms, Transparent Screens, and 3D Laser Projections (2003), `http://www.laser-magic.com/transscreen.html`
11. Sparacino, F., Davenport, G., Pentland, A.: Media in Performance: Interactive Spaces for Dance, Theater, Circus, and Museum Exhibits. IBM Syst. J. 39(3), 479–510 (2000)
12. Wei, S., Fortin, M., Navab, N., Sutton, T.: Ozone: Continuous State-based Media Choreography System for Live Performance. In: ACM Multimedia, pp. 1383–1392. ACM Press (2010)