

Empathetic Interactive Music Video Experience

Myunghee Lee
Digital Experience Lab.
Korea University
revise1@naver.com

Youngsun Kim
Digital Experience Lab.
Korea University
zyoko85@korea.ac.kr

Gerard Jounghyun Kim
Digital Experience Lab.
Korea University
gjkim@korea.ac.kr

ABSTRACT

Empatheater is a video playing system that is controlled by multimodal interaction. As the video is played, the user must interact and emulate predefined “events” for the video to continue on. The user is given the illusion of playing an active role in the unraveling video content and can empathize with the performer. In this paper, we report about user experiences with Empatheater when applied to musical video contents.

Keywords

Music video, Empathy, Interactive video, Musical event, Multimodal interaction.

1. INTRODUCTION

Aside from lone audio, music video is another popular format for appreciating music. In [6], we proposed an interactive video system, named “Empatheater.” In Empatheater, as the video is played, the user must interact and emulate predefined “events” (authored beforehand) through appropriate multimodal and whole body interaction. Without the timely interaction, the video stops, and the system shows additional information as how to continue. Empatheater was designed with the emphasis on how to lead the user to feel empathy e.g. with main content driving character (see Figure 1).

In this paper, we have applied Empatheater to appreciating music empathetically through interaction with music videos. Music listeners often imagine themselves being the singer, dancer, conductor and mimic them in various ways. We believe that Empatheater can promote and upgrade such experiences even further and offer a more active way of enjoying music. We applied Empatheater for a small user group to experience four different types of music videos and present findings in terms of their user experiences and modifications needed.

2. RELATED WORK

Interactive video was first proposed by Lippman in their seminal Aspen Movie Map system [1]. However, at the time, the focus was more on the interactivity itself, offering different endings through interaction and non-linear story structure. Recently, interesting works on non-conventional interaction with video contents have been proposed. For example, Hurst has proposed for various forms of touch screen and gesture based interface for video browsing [2][3][4]. However, most of

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
NIME'12, May 21-23, 2012, University of Michigan, Ann Arbor.
Copyright remains with the author(s).

these works were geared toward the problem of efficient browsing rather than transforming the video content itself into user experience.



(a)



(b)

Figure 1: (a) Screen capture of Empatheater. Musical events to interact to appear in the bottom time line. (b) User interacting using a Wii-mote.

Jehan et al. presented a work in which the pace of a “dancing” video was controlled by different musical beats [5]. The altered pacing (according to various beats) attributed to eliciting different experiences from a same video clip. Empatheater extends this idea by introducing more interactivity and empathy. Empatheater also bears some similarity to the music rhythm games where the user must interact timely to musical events. However, they fall short of creating any empathy due to the lack of controllability of the musical flow [7].

3. EMPATHEATER

Empatheater is implemented as an ordinary video player but with an added capability to interpret interaction information. In order to interact with the video (and stop or forego the playing), interaction events must be specified and stored beforehand. Event information amounts to the type of the interaction and point at which (time) it happens. For this, we have also implemented a simple authoring tool through which the aforementioned information can be specified with ease. The type of required/desired interaction (both input and output) is

also specified by the author (e.g. “Space bar (keyboard),” “acceleration sensor with value $> x$,” “play sound file”, etc.).

In order to help user interact with the video (during play time) at the right time in the right way, a basic timeline/bar widget is shown and in addition iconic cues are displayed to foretell the type of upcoming interaction. However, when the user fails to interact in the right way (e.g. no or “weak” interaction), the video stops and further hints are displayed. Similarly to the interaction points themselves, these guidance features are inserted manually using the aforementioned authoring tool.

We believe in principle that the interaction must be multimodal, whole body based and inspired from the content of the video, to elicit the empathetic feeling. As such, musical events are to be enacted with “musical” interaction, e.g. MIDI keyboard strike, Wii-mote action motion, voice activation, etc. The author can freely design such interaction (both input and output) and store it along with the event information to be interpreted during the play. Indication of successful interaction is also important in giving the illusion of control to the user. In the middle of Figure 1(a) is a graphic firework in response to a timely interaction (e.g. dance motion, conducting gesture, voice activation) by the user.

4. USER STUDY

We chose four different types of music videos to investigate the users’ empathetic musical experience: (1) interacting in motion (simple tapping) to (and enacting to perform) relatively slow melodic events (excerpt from Mozart Piano Concerto No. 21, 2nd movement), (2) interacting vocally to medium tempo melodic events (Beethoven’s Ode to Joy), (3) interacting in (dance-like) motion to fast tempo but mostly regular beats (Korean pop dance song), and (4) interacting in (conducting-like) motion to accentuated or pompous musical events (Beethoven’s 3rd Symphony, final movement).

In this pilot experiment, we asked twelve subjects (college students, average age: 22.5) to experience the four above music videos by (1) mere passive listening and (2) interacting using Empatheater (one factor, within-subject, repeated measure). The experiment proceeded as follows. In condition 1 (“Passive”), the user simply sat, watched and listened to the short music video. In the second condition (“Empatheater”), the user used various methods (e.g. tap, voice, swing motion, and conduct like gestures) to enact the musical events (marked in the time line of Empatheater, see Figure 1) in the respective video. The orders of the types of the music video given and interaction conditions were all counter-balanced. The subject completed a short survey asking about user experience mainly through the level of empathy and presence (answered in 5 Likert scale: 0~4), and preference (tabulated in percentages).

Figure 2 shows the major experimental results (due to shortage of space, we only show responses to four survey questions). Users answered with Empatheater they experienced much higher empathy with the main character (e.g. performer in the video) and felt like actually carrying out the musical task (Figures 2(a) and (b)). They answered they were able to understand the musical content more deeply (Figure 2(c)). These results were all with statistically significant differences to the “Passive” condition. Although not shown a similar results was obtained regarding the level of immersion and presence felt. The users also reported among the four types musical videos, “Conducting” was preferred the most (Figure 2(d)) and was also identified as the most effective content type for Empatheater. The “Dance” was felt most ineffective and

least preferred because the subjects did not feel comfortable even in the experimental setting to openly dance and because the motion gestures were limited to just simple Wii-mote swings.

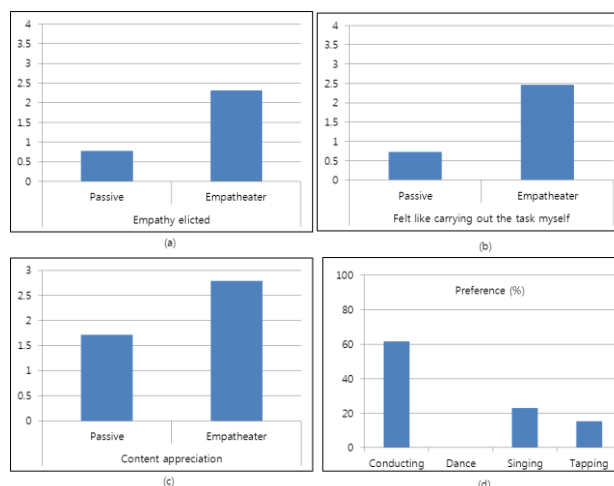


Figure 2: Experimental results: Passive (left) vs. Empatheater (right) (a) Empathy elicited, (b) How much felt as carrying out the task (e.g. conducting, playing the piano, singing, dancing), (c) How much the user appreciated the given content and (d) Preference among the music video types (in percentages).

5. CONCLUSION

In this paper, we have introduced “Empatheater,” a video system activated by multimodal interaction. Our pilot study has shown that when applied to music videos, it can help user identify with the performer and help comprehend the music more deeply. We believe Empatheater can provide a new way of experiencing and appreciating music, in an “active” way.

6. ACKNOWLEDGEMENTS

This research was supported by the Ministry of Knowledge Economy and KIAT through the “Workforce Development Program in Strategic Technology”.

7. REFERENCES

- [1] Lippman, A. Movie-maps: An application of the optical videodisc to computer graphics. *Proc. of the 7th annual conference on Computer graphics and interactive techniques*, 1980, pp. 32–42.
- [2] Hurst, W. Interactive audio-video browsing. *Proc. of ACM Multimedia Conference*, 2006.
- [3] Dragicevic, P., Ramos, G., Bibliowicz, J., Nowrouzezahrai, D., Balakrishnan, R. and Singh, K. Video browsing by direct manipulation. *Proc. of ACM CHI*, 2008.
- [4] Goldman, D., Gonterman, C., Curless, B., Salesin, D. and Seitz, S. Video object annotation, navigation and composition. *Proc. of UIST*, 2008.
- [5] Jehan, T., Lew, M. and Vaucelle, C. Cati dance: self-edited, self-synchronized music video. *SIGGRAPH Technical Sketch, Conference Abstracts and Applications*, 2003.
- [6] Lee, M., Kim, G. J. Empathetic video experience through timely multimodal interaction. *ICMI-MLMI*, 2010.
- [7] Wikipedia, BeatMania. <http://en.wikipedia.org/wiki/Beatmania>, 2012