Measurement of Latency in Interactive Multimedia Art

Yoichi Nagashima

Shizuoka University of Art and Culture / Art & Science Laboratory
1794-1 Noguchi-cho, Hamamatsu, Shizuoka 430-8533, Japan

nagasm@computer.org

Abstract

In this paper, I would like to introduce my experimental study of multimedia psychology. My initial focus of investigation is the interaction between perceptions of auditory and visual beats. When the musical and graphical beats are completely synchronized with each other, as in a music video for promotional purposes, the audience feels that they are natural and comforting. My initial experiment has proved that the actual tempos of music and images are a little different. If a slight time lag exists between the musical and pictorial beats, the audience tries to keep them in synchronization by unconsciously changing the interpretation of the time-based beat points. As the lag increases over time, the audience seems to perceive that the beat synchronization has changed from being more downbeat to more upbeat, and continues enjoying it. I have developed an experiment system that can generate and control out-of-phase visual and auditory beats in real time, and have tested many subjects with it. This paper describes the measurement of time lags generated in the experiment system, as part of my psychological experiment.

1. Introduction

Murao investigated how people perceived musical beats and defined three types of beats as follows:

The "tago" beat is particular to the Japanese people, who perceive that it strikes and stops at its trailing edge (see Fig. 1).

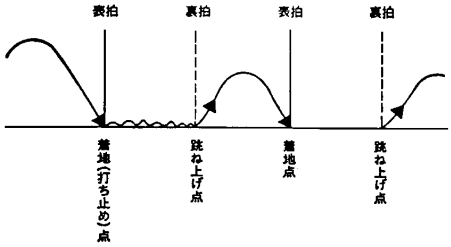

Figure 1 : Tago Beat

The "toe" beat is particular to Europeans and Americans, who perceive that it rebounds instantaneously at its trailing edge (see Fig. 2).

The "heel" beat is pervasive in rock'n roll and pop music, and the audience perceives that it linearly advances toward the beat point, rather than rebounding (see Fig. 3). This research involved using only the heel beat to examine the relationship between visual and auditory beats.

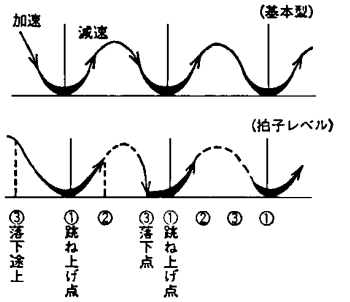

Figure 2 :Toe Beat

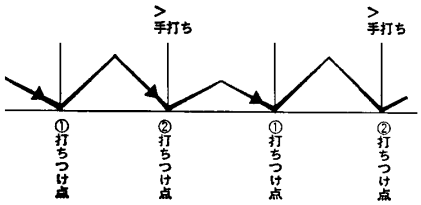

Figure 3 : Heel Beat

2. Model of "Drawing-in Effect "

In 2002, I created a film titled "JizoGazoDazo." This film consists of a piece of 5-beat background music and images transformed and connected at regular intervals by means of morphing. The latter serves as a visual beat. Because the tempo of the music deviates slightly from the constant intervals of the moving images - the pictorial beat, the matching points of the musical and pictorial beats gradually changes over time, which is indicated graphically in Fig. 4. The audience unconsciously selects the nearest musical beat to the pictorial beat, therefore the musical beat seems to respond to the pictorial beat and to alternate between the downbeat and upbeat. I call this phenomenon the "drawing-in effect." Fig. 5 illustrates the psychological experiment system that is designed to generate time lags between the visual and auditory beats. I instructed a test subject who watched a film while listening to continuously playing music, to tap a switch at the points where he/she perceived each pictorial beat.

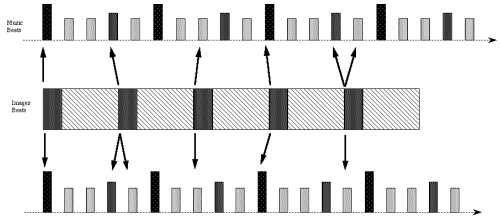

Figure 4 : Beat shifts in "JizoGazoDazo"

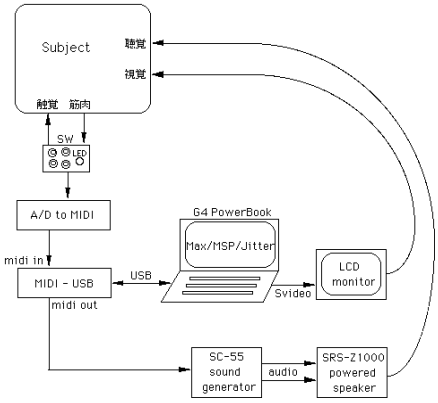

Figure 5 : System of the psycological experiments

3. Psycological Experiment System

I constructed this experiment system using the application Max/MSP/Jitter. Fig. 6 shows the main patch of the experiment system. This paper does not describe how to compose the pictorial and musical elements used in the experiment due to limitations of space. For detailed information about the composition, see References 2 and 3.

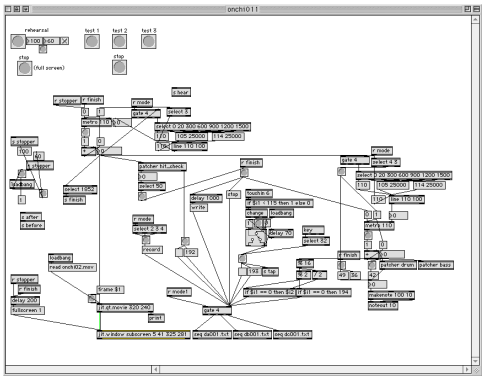

Figure 6 : Main pateh of the system

The test subject watched the film on an LCD monitor placed at his/her front. Simultaneously, he/she relaxed and listened to the continuously playing music from the stereo monitor. I instructed the subject to tap a chosen switch out of four the moment each picture specified as a visual beat appeared. The testing process consisted of the following four steps: (1) Rehearsal for getting accustomed to tapping the switch; (2) Experiment Mode 1, in which the subject taps the switch under the circumstances that the pictorial beats are indicated by a guiding sound; (3) Experiment Mode 2, in which the actual moving images appear without the guiding sound, and; (4) Experiment Mode 3, consisting of an alternate experiment in which the subject undergoes the same test as in Experiment Mode 2. Fig. 7 includes a photo showing the experiment environment.

Figure 7 : Photo example of the experiments

Each of Experiment Modes 2 and 3 requires about three minutes to complete. In Experiment Mode 2, the interval between pictorial beats is precisely constant, but that between musical beats is slightly lengthened or shortened alternately. In Experiment Mode 3, the interval between musical beats is kept constant, but that between pictorial beats varies slightly.

4. Measurement of System Latency

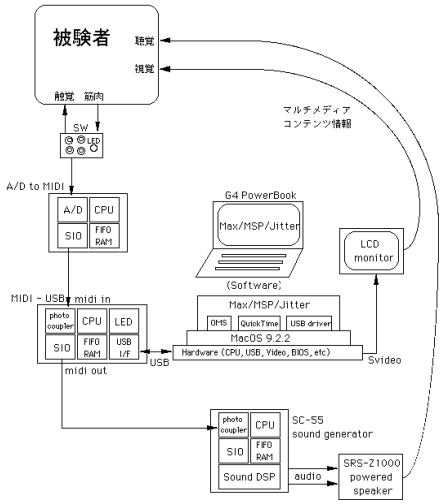

Figure 8 : Detail of the System

I carefully analyzed the system (see Fig. 8) and designed eleven experiments to measure the system delays. Of the eleven experiments, only three are described in this paper due to limitations of space. I will provide the details and results of all of the experiments in my poster session.

4-1. Experiment 1

Fig. 9 presents the details of Experiment 1. Trigger 1 is

activated by the signal from the small switch tapped by the subject the moment each pictorial beat appears. Trigger 2 is activated by the signal from the photocoupler PC910 of the Roland MIDI-USB interface UM-2. In this experiment, I measured the combined delay of the A/D converter, the software running on the CPU, the FIFO buffer of the internal MIDI, and the MIDI output circuit, which is referred to as the A/D-to-MIDI delay. Fig. 10 shows one of the 20 results of Experiment 1. In this experiment, the delay from the triggered switch to the MIDI input of PC910 is about 14 milliseconds.

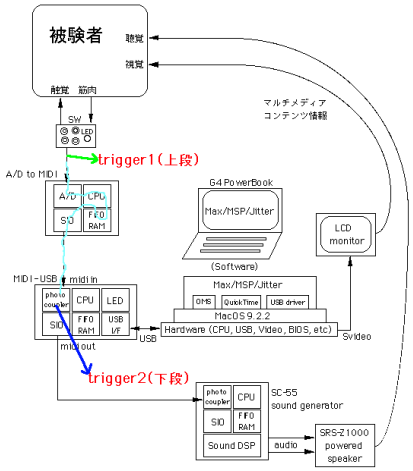

Figure 9 : Experiment 1

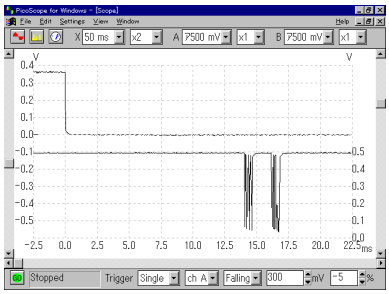

Figure 10 : One of the results of Experiment 1

4-2. Experiment 3

Fig. 11 presents the details of Experiment 3. Trigger 1 is activated by the input signal from PC910, and Trigger 2 is activated by the output signal from the Roland sound generator SC-55. Signals generated by tapping the switch pass through UM-2 and are processed by the application Max/MSP/Jitter running on a Macintosh computer connected to the MIDI-USB converter of UM-2. The resulting note information is fed to SC-55 via UM-2 and the corresponding sound is generated.

Fig. 12 shows one of the 20 results of Experiment 3. In this experiment, the delay from the switch tapped to the sound generated is about 13 milliseconds. This lag is shorter than the delay of 14 milliseconds resulted from Experiment 1, which was due to deficiencies in each test that were greater than the differences in those in the latter experiment.

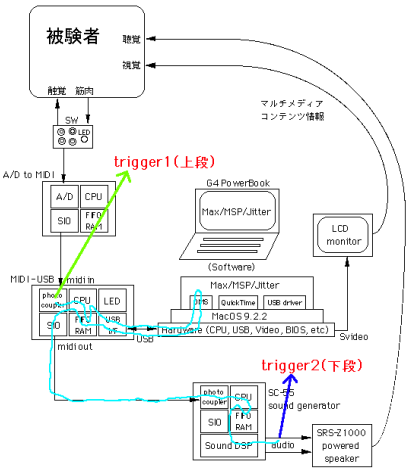

Figure 11 : Experiment 3

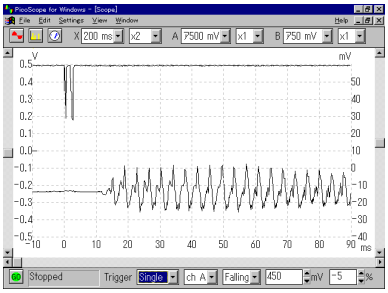

Figure 12 : One of the results of Experiment 3

4-3. Experiment 9

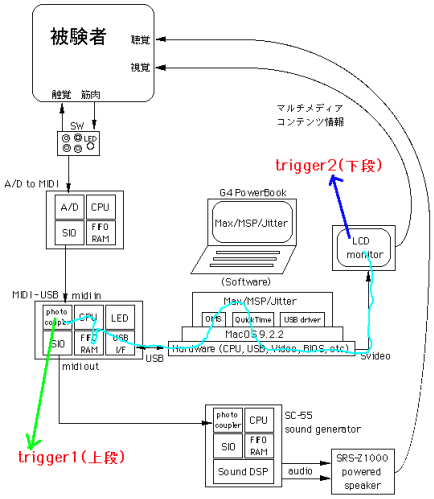

Figure 13 : Experiment 9

Fig. 13 presents the details of Experiment 9. Trigger 1 is activated by the input signal from PC910, and Trigger 2 is activated by the signal from the high-speed

photosensor (SHARP 1S474) affixed to the screen of the LCD monitor connected to the video mirroring output of the Macintosh computer. The system uses Max/MSP/Jitter to change the brightness of the screen from darkest to the maximum brightness according to the MIDI input signal generated by tapping the switch.
Fig. 14 shows one of the 20 results of Experiment 9. In this experiment, the delay from the triggered switch to the screen being fully illuminated is about 60 milliseconds.

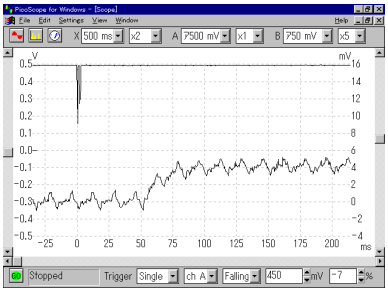

Figure 14 : One of the results of Experiment 9

4-4. Experiments Results

Fig. 15 indicates the 20 results for each of the eleven experiments. Some data concentrate in a narrow area, but some data are spread widely. It is very important to understand this kind of data variation, otherwise there is the risk of making the results of psychological experiments meaningless.

実験データ(実測値) : 単位 = msec

	実験 1	実験 2	実験 3	実験 4	実験 5	実験 6	実験 7	実験 8	実験 9	実験 10	実験 11
data 1	18.0	6.0	12.0	10.0	3.5	9.0	13.0	16.0	75	5.0	0.15
2	11.0	6.0	10.0	9.0	4.0	11.0	8.0	13.0	30	4.0	0.15
3	10.0	4.5	10.0	11.0	6.0	8.0	10.0	14.0	35	5.0	0.15
4	8.0	4.0	8.0	8.0	4.5	12.0	11.0	18.0	30	3.0	0.15
5	14.0	7.0	8.0	10.0	5.5	12.0	10.0	13.0	40	7.0	0.15
6	10.0	6.0	9.0	9.0	6.0	8.0	12.0	15.0	55	5.0	0.15
7	3.0	3.5	11.0	8.0	7.0	12.0	11.0	13.0	35	4.5	0.15
8	16.0	5.0	8.0	11.0	4.0	8.0	9.0	15.0	30	5.5	0.15
9	13.0	4.0	10.0	10.0	5.0	13.0	12.0	16.0	50	7.5	0.15
10	18.0	5.5	8.0	11.0	4.0	10.0	9.0	15.0	40	5.5	0.15
11	6.0	5.5	11.0	11.0	7.0	11.0	10.0	17.0	45	7.0	0.15
12	10.0	3.0	11.0	8.0	4.0	13.0	8.0	17.0	25	7.5	0.15
13	7.0	5.0	9.0	8.0	8.0	11.0	9.0	16.0	100	6.5	0.15
14	5.0	5.0	12.0	10.0	7.0	12.0	9.0	15.0	40	3.5	0.15
15	16.0	5.0	10.0	12.0	4.0	12.0	12.0	14.0	25	7.0	0.15
16	13.0	4.0	9.0	9.0	4.5	9.0	12.0	17.0	120	7.0	0.15
17	6.0	5.5	10.0	11.0	6.5	12.0	8.0	13.0	40	8.0	0.15
18	9.0	5.0	10.0	9.0	5.0	8.0	10.0	13.0	70	5.0	0.15
19	15.0	4.5	11.0	9.0	8.0	9.0	12.0	14.0	25	7.0	0.15
data 20	7.0	7.5	9.0	10.0	7.0	11.0	11.0	16.0	35	3.5	0.15
平均	10.8	5.1	9.8	9.7	5.5	10.6	10.3	15.0	47.2	5.7	0.15
最大値	18.0	7.5	12.0	12.0	8.0	13.0	13.0	18.0	120	8.0	0.15
最小値	3.0	3.0	8.0	8.0	3.5	8.0	8.0	13.0	25	3.0	0.15
標準偏差	4.459	1.115	1.281	1.218	1.464	1.761	1.559	1.589	25.62	1.525	0.00

Figure 15 : Results of all Experiments

Fig. 16 shows an example of the Max/MSP/Jitter window in which the results from the psychological experiment of my initial research appears. The green, yellow and blue dots indicate pictorial beats, tapping timing generated by subjects perceiving them, and musical beats respectively. Moreover, the sky blue dots correspond to accents. This window allows the data to be enlarged, rotated and translated in a 3-D space.
Fig. 17 presents the enlarged test results for a certain subject. Though the interval between pictorial beats is constant, the dots corresponding to visual tapping timing provided by the subject appear in the vicinity of musical beats having a slightly different tempo from the pictorial ones. My hypothesis is that this is precisely the "drawing-in effect."

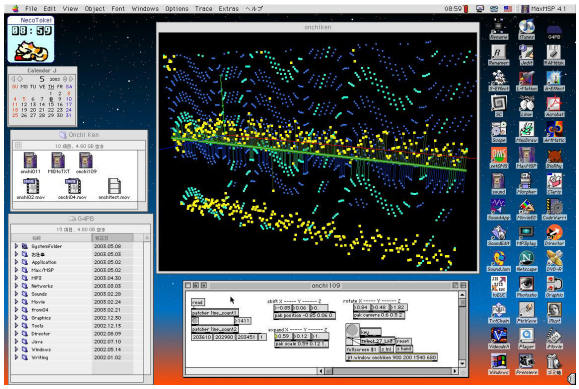

Figure 16 : Sample of the Result Data Visualize

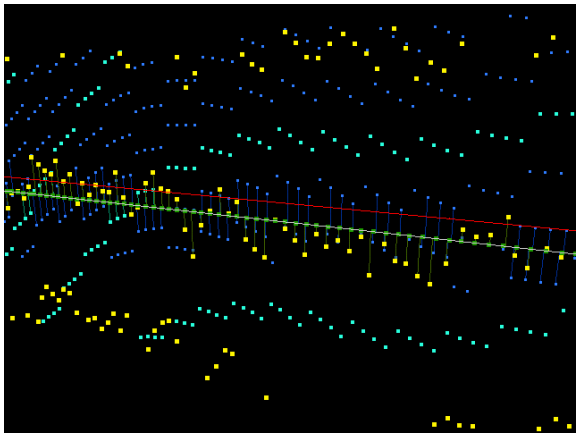

Figure 17 : One data showing "Drawing-in Effect"

5. Conclusions

I have introduced some studies of experimental applications and delay measurements for the interaction between human and computer in multimedia information processing. Interactive multimedia arts comprise an interesting and experimental field for the research of human interfaces or perception. I will further investigate this topic in conjunction with the conducting of many experiments.

6. References

[1] Tadahiro Murao, *Discovery from "Tago Rhythm" (in Japanese)*, Research of Musical Education 56, Ongaku-no-Tomo, pp.177-190, 1988

[2] Yoichi Nagashima, *Drawing-in Effect on Perception of Beats in Multimedia (in Japanese)*, http://suac.net/nagasm/ASL/beat/

[3] Yoichi Nagashima, *Art and Science Laboratory (in Japanese)*, http://nagasm.org