Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

Congruency between music and motion pictures in the context of video games: Effects of emotional features in music

Shinya Kanamori,[*1] Ryo Yoneda,[*2] Masashi Yamada[*3]

[*] *Graduate School of Engineering, Kanazawa Institute of Technology, Japan*
[1] s_kanamori@venus.kanazawa-it.ac.jp, [2] r_yoneda@venus.kanazawa-it.ac.jp,
[3] m-yamada@neptune.kanazawa-it.ac.jp

ABSTRACT

In the present study, two experiments are conducted. In the first experiment, it is revealed that the impression of game music is spanned by "pleasantness" and "excitation" axes, using one hundred pieces of game music. In the second experiment, it is shown that the congruency of moving picture and musical tune does not decrease and the whole impression is not change significantly, even if a tune is replaced by a tune which possesses similar impression. These results suggests that an archive, where various tunes are plotted on the impression plane spanned by the "pleasantness" and "excitation" axes, is useful to communicate in the group of game creators and engineers, for designating a piece of music for a scene in a video game.

I. INTRODUCTION

In recent years, the cost, time and the amount of human resources for creating video game content is significantly increasing, because the hardware of consumer game machines and PCs is increasing greatly in their quality. However even now game creators create contents with a trial and error process based on their experience and intuitive belief. This is true in the situation of composing and arranging a piece of game music. Several musicians and engineers join the team for creating musical pieces for a game title. They have to share their image for pieces of music, and communicate their image with visual designers, system designers, directors and engineers. For such communication and sharing of musical images, a scientific basis for designating a piece of game music is strongly needed.

In the present study, we conducted two experiments to obtain a scientific basis of designing pieces of game music. In the first experiment, dimensions which construct the impression of game music were revealed. In the second experiment, the congruency between music and moving picture is revealed in the context of scenes in video games. These results show a scientific basis for the communication in the group of game creators and engineers, for designating a piece of music for a scene in a video game.

II. EXPERIMENT 1

A. Method

One hundred pieces of game music was selected from commercially produced sound track CDs of video games. These pieces are from various genres of videogames, *e.g.*, roll playing games, action games, survival horror games, *etc.* Fifteen students from the Kanazawa Institute of Technology participated the experiment as listeners. They were requested to listen to each of the 100 pieces through the STAX Lambda-pro headphones at the level of LAeq= 62~82 dB. After listening to each piece, the listeners rated their impression of the piece using 24 semantic differential (SD) scales listed in Table 1. Each of the 24 scale was a seven-step bipolar scale. The listeners took a rest period more than ten minutes after each ten pieces. Each listener completed the experiment during two to three days.

B. Results and Discussion

The scores of a SD scale for a piece were averaged over the listeners. Then, the averaged scores were used for factor analysis (principal factor analysis). Results of the analysis showed that the two-factor solution accounted for 79 % of the data variance. Table 1 shows the factor loadings of each SD scale on the two factors. The factors are labeled "pleasantness" and "excitation", respectively.

Russell (1980) showed that the emotion is illustrated by a circumplex model, which is spanned by "valence" and "activity" dimensions. Hevner (1936, 1937) plotted eight clusters of adjectives in circles. The plots of the adjectives were consistent with the Russell's circumplex model. The two dimensional plane the results of the present experiment shows is consistent to the Russell's circumplex model and Hevener's adjective circle. And the factors of "pleasantness" and "excitation" are correspond to Russell's "valence" and "activity" dimensions, respectively.

Each of the piece is plotted on the two-dimensional plane using the factor scores (Fig. 1). Using these plots, a director may covey the image of a musical piece with musicians as follows: "Please compose a tune with the impression like these pieces for this picture." The team of musicians and engineers also can share the image of a musical piece as follows, "For this scene, an active but unpleasant tune is suitable."

However, in the context of video games, the image communication using the system described above does not work effectively, if a musical tune is congruent with a moving picture but a different musical tune is not congruent, even the two tunes possess similar impression.

Bolivar, et al. (1994) pointed out that congruency between auditory and visual stimuli can be divided into semantic and formal congruency. Iwamiya and his colleagues has been conducting a series of studies on both congruencies, using computer graphics, subtitles, music, sound effects, *etc*. as materials for the audio-visual stimuli (e.g., Iwamiya, 1994; Iwamiya *et. al*, 2002, Kim, Iwamiya and Kitano, 2009). Recent video games are one type of content which effectively uses the congruency between auditory and visual stimuli.

In Experiment 2, the semantic congruency between moving pictures and musical pieces are investigated, in the context of video games.

III. EXPERIMENT 2

A. Method

One hundred musical pieces used in Experiment 1 are plotted in Fig. 1. Four pieces from the 100 pieces were selected for the present experiment. Each piece was from each quadrant in Fig. 1. These four pieces are called "original tunes" and noted as Music 1 - Music 4, in the present study. Then, the four moving pictures which appeared in video games with the original tunes were prepared. These moving pictures are noted as Scene 1 - Scene 4. These moving pictures included sound effects. We also selected the other four musical tunes. Each of them was from each quadrant in Fig.1. They are noted as Music 1' – Music 4'. Music1', 2', 3' and 4'posseessed similar impression to Music 1, 2, 3 and 4, respectively. These musical pieces are listed in Table 2.

For each moving picture, one original tune, one similar-impression tune, three different-impression tunes, and no music were combined and six audio-visual stimuli were constructed. For example, Scene 1 was combined with Music 1, Music 1', Music 2, Music 3, Music 4 and No Music. In total, 24 stimuli listed in Table 3 were constructed for the present experiment.

Ten students from the Kanazawa Institute of Technology participated in the experiment. The participants were presented each of the 24 stimuli and then requested to rate the impression of it, using 24 seven-step SD scales shown in Table 4. Except for No Music conditions, the participants were requested to rate the degree of the congruency between the musical tune and moving picture in a seven-step scale, and describe what the character in the scene was doing or was going to do, freely.

The moving pictures were projected on a large screen and the musical pieces and sound effects were presented through the STAX Lambda-pro headphones at the level of LAeq= 62~83 dB. After completing every twelve trials, the subjects took a ten-minute rest interval.

B. Results and Discussion

The scores of a SD scale for a piece were averaged over the participants. Then, the averaged scores were used for factor analysis. Results of the analysis showed that the two-factor solution accounted for 78 % of the data variance. As shown in Table 4, the two factors were labeled "pleasantness" and "excitation", again.

Figure 2 shows the plots of audio-visual stimuli on the "pleasantness" and "excitation" plane. In Fig. 2, the same or similar-impression musical tunes cluster in a small area. This implies that the impression of the audio-visual stimuli is almost determined by music rather than moving picture. These results were consistent to the results of Yamada (2008) in which a survival horror game was used. Figure 2 shows that the impression of a scene with the original tune is very near the scene with the similar-impression tune, e. g., Scene 1 [Music 1] is near Scene 1[Music 1']. This implies a musical piece which was composed for a scene can be replaced by a similar-impression tune without change the whole impression of the audio-visual stimuli.

Figure 2 also shows that three stimuli with No Music are rated as unpleasant and unexcited. Scene 4 without music is rated as pleasant because the scene included wave sound and bird twittering.

The degree of the congruency between the musical tune and moving picture were averaged over the participants for 16 stimuli where musical pieces are combined with moving pictures. The 16 stimuli are plotted on the scale of the congruency in Fig. 3. In Fig. 3, a stimulus where a motion picture with an original musical tune shows a high value of the congruency. Figure 3 also shows that a stimulus where a similar-impression tune replaced the original tune does not decrease the congruency, significantly. The results of the free description also showed that the interpretation of the behavior of the character in the scene, in principle, does not change by replacing a original tunes with a similar-impression tune. For Scene 2, the congruency is higher for Music 2' than Music 2. This implies that the game producer of Scene 2 may admit of replace more suitable tune for the scene.

The congruency most of the stimulus where a different-impression tune replaced the original tune shows a low degree of congruency. Even the case the congruency shows a high value, the interpretation for the behavior of the character of the scene drastically changed from the scene with the original tune as follows: "The character running to another stage" changes to "the character is going to start a new battle."

IV. CONCLUSION

In the present study, it is revealed that the impression of game music is spanned by "pleasantness" and "excitation" axes. Moreover, it is shown that the congruency of moving picture and musical tune does not decrease and the whole impression is not change significantly, even if a tune is replaced by a tune which possesses similar impression. These results suggests that an archive, where various tunes are plotted on the impression plane spanned by the "pleasantness" and "excitation" axes, is useful to communicate in the group of game creators and engineers, for designating a piece of music for a scene in a video game.

ACKNOWLEDGMENT

This study is partly supported by KAKEN (22615043).

REFERENCES

Bolivar, V. J., Cohen, A. J., and Fentress, J. C. (1994) Semantic and formal congruency in music and motion pictures: Effects on the interpretation of visual action, *Psychomusicology*, *13*, 28-59

Heavner, K. (1935). Expression in music: A discussion of experimental studies and theories, *Psychological Reviews*, **42**, 186-204.

Heavner, K. (1937). The affective value of pitch and tempo in music, *American Journal of Psychology*, **48**, 621-630.

Iwamiya, S. (1994). Interactions between auditory and visual processing when listening to music and audio visual context: 1. Matching 2. Audio quality, *Psychomusicology*, **13**, 133-154.

Iwamiya, S., Jogetsu, Y., Sugano, Y. and Takada M. (2002). Effects of musical tempo and tonality and visual speed and density on perception of audiovisual content, *Journal of Music Perception and Cognition*, **8**, 53-64.

Kim, K. Iwamiya, S. and Kitano H. (2009) . Subjective Congruence between rotating movement of a visual image and pitch

modulation of a sound, *Japanese Journal of Ergonomics*, **45**, 336-343.

Russell, J. A. (1980). A circumplex model of affect, *Journal of Personality and Social Psychology*, 3**9**, 1161-1178.

Yamada, M. (2008). The effect of music on the fear emotion in the context of a survival-horror video game, *Proc. 10th International Conference on Music Perception and Cognition*, 594 – 597.

Yoneda, R. and Yamada, M. (2009). A multi-dimensional study on the emotion in popular music listened to by young Japanese listeners, *Proc. 10th Western Pacific Acoustics Conference*, 216

Table 1 Semantic differential scales used in Experiment 1 and factor loadings of them

SD Scale		Factor	
		Pleasantness	Excitation
Dark	– Bright	0.96	0.00
Heavy	– Light	0.87	-0.27
Tense	– Relaxed	0.70	-0.66
Pleasant	– Unpleasant	-0.95	0.13
Warm	– Cold	-0.90	0.09
Clear	– Unclesr	-0.90	0.13
Cute	– Uncute	-0.89	0.34
Fresh	– Sordid	-0.89	0.32
Delightful	– Dull	-0.87	-0.33
Cheerful	– Gloomy	-0.80	-0.12
Clean	– Dirty	-0.77	0.32
Weak	– Strong	-0.26	0.86
Tranquil	– Restless	-0.28	0.85
Loose	– Tight	-0.30	0.83
Monotonous	– Varied	0.25	0.58
Excited	– Unexcited	0.10	-0.95
Showy	– Humble	-0.20	-0.90
Agitated	– Calm	0.12	-0.90
Pwerful	– Powerless	0.29	-0.88
Speedy	– Slowly	-0.07	-0.86
Hard	– Soft	0.55	-0.76
Mixed	– Neat	0.48	-0.72
Wide	– Narrow	0.16	-0.68
Impressive	– Unimpressive	-0.42	-0.53
Contribution Rate		0.393	0.389

Table 2 Musical pieces used in Experiment 2

No.	Music Name	Game Title	Note
1	Honnoji-dera (Temple)	Sengoku Musou 2 (Samurai Warriors 2)	Music 1
2	ED-ZOMBIE	biohazard 0	Music 2
3	Stage IV	Devil May Cry 4	Music 3
4	Umi no ue no mura, Moga (Prince of Moga)	Monster Hunter 3	Music 4
5	For Achieve	STAROCEA&VALKYRIE	Music 1'
6	Ganado II	biohazard 4	Music 2'
7	Shiren no Aida (The Trials)	Final Fantasy X	Music 3'
8	Mikan jii no toujou (Mr. Orange Appears)	Okami	Music 4'

Table 3 Audio-visual stimuli used in Experiment 2

No.	Note	Game Title of Moving Picture	Title of Musical Piece
1	Scene 1 [No music]	Samuarai Warrior 2	None
2	Scene 1 [Music 1']	Samuarai Warrior 2	For Achieve
3	Scene 1 [Music 1]	Samuarai Warrior 2	Honno Temple
4	Scene 1 [Music 2]	Samuarai Warrior 2	ED-Zombie
5	Scene 1 [Music 3]	Samuarai Warrior 2	Stage IV
6	Scene 1 [Music 4]	Samuarai Warrior 2	Prince of Moga
7	Scene 2 [No music]	biohazard 0	None
8	Scene 2 [Music 2']	biohazard 0	Ganado II
9	Scene 2 [Music 1]	biohazard 0	Honno Temple
10	Scene 2 [Music 2]	biohazard 0	ED-Zombie
11	Scene 2 [Music 3]	biohazard 0	Stage IV
12	Scene 2 [Music 4]	biohazard 0	Prince of Moga
13	Scene 3 [No music]	Devil May Cry 4	None
14	Scene 3 [Music 3']	Devil May Cry 4	The Trials
15	Scene 3 [Music 1]	Devil May Cry 4	Honno Temple
16	Scene 3 [Music 2]	Devil May Cry 4	ED-Zombie
17	Scene 3 [Music 3]	Devil May Cry 4	Stage IV
18	Scene 3 [Music 4]	Devil May Cry 4	Prince of Moga
19	Scene 4 [No music]	Monster Hunter 3	None
20	Scene 4 [Music 4']	Monster Hunter 3	Mr. Orange Appears
21	Scene 4 [Music 1]	Monster Hunter 3	Honno Temple
22	Scene 4 [Music 2]	Monster Hunter 3	ED-Zombie
23	Scene 4 [Music 3]	Monster Hunter 3	Stage IV
24	Scene 4 [Music 4]	Monster Hunter 3	Prince of Moga

Table 4 Semantic differential scales in Experiment 2 and factor loadings of them

SD Scale		Factor	
		Pleasantness	Excitation
Cheerful	– Gloomy	-0.95	-0.12
Pleasant	– Unpleasant	-0.94	0.14
Warm	– Cold	-0.93	0.14
Clear	– Unclear	-0.90	-0.05
Adorable	– Detestable	-0.85	0.39
Fresh	– Shabby	-0.84	0.34
Loose	– Tight	-0.72	0.60
Clean	– Dirty	-0.72	0.41
Dark	– Bright	0.96	0.00
Heavy	– Light	0.92	-0.14
Hard	– Soft	0.87	-0.36
Tense	– Relaxed	0.84	-0.51
Showy	– Humble	-0.15	-0.94
Excited	– Unexcited	0.30	-0.89
Agitated	– Calm	0.25	-0.87
Powerfull	– Pawerfullness	0.38	-0.87
Speedy	– Slowly	0.93	-0.78
Wide scale	– Narrow scale	0.06	-0.76
Impressive	– Unimpressive	0.16	-0.71
Weak	– Strong	-0.40	0.87
Tranquil	– Restless	-0.50	0.75
Monotonous	– Varied	-0.43	0.18
Mixed	– Neat	0.25	-0.69
Delightful	– Dull	-0.46	-0.70
Contribution Rate		0.43	0.35

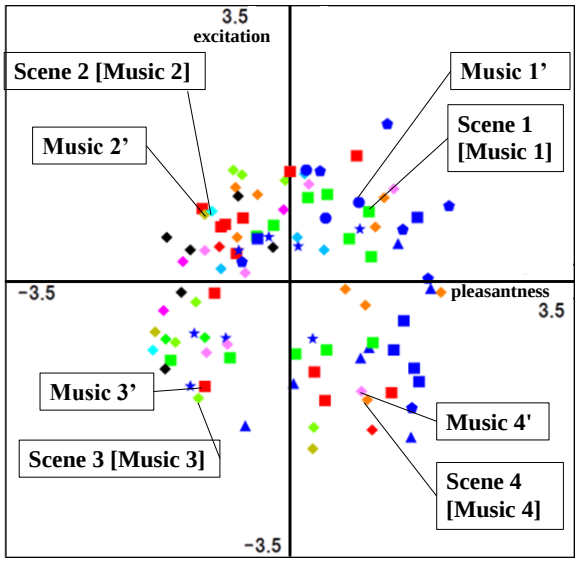

Figure 1. Impression for pieces of game music

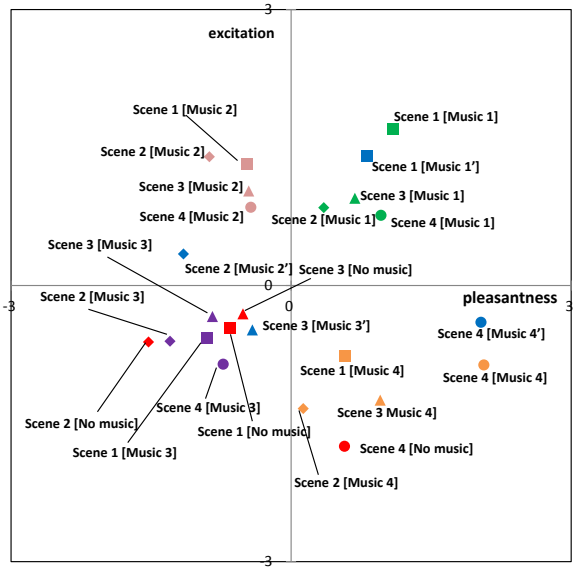

Figure 2. Impression of audio-visual stimuli used in Experiment 2

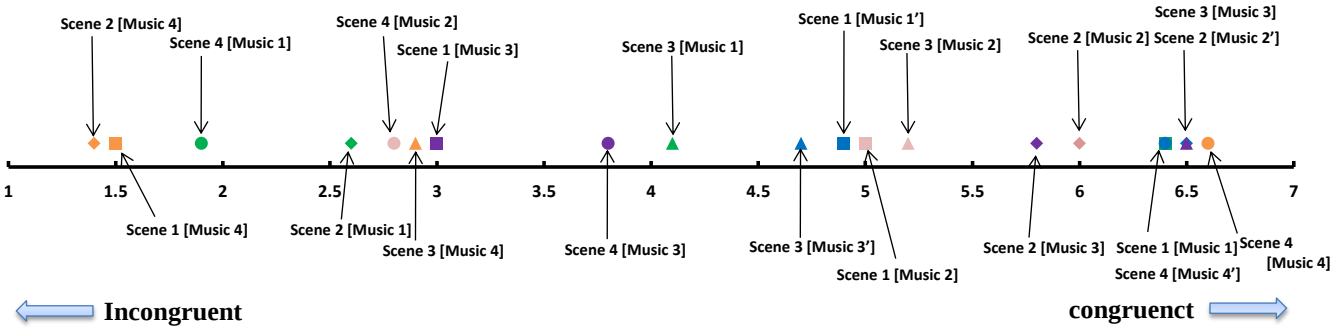

Figure 3. Congruency between musical tune and scenes