비쥬얼뮤직의 정서 감정: 추상적 멀티미디어 정보 전달을 위한 두개의 조합모델 검증

The Perception of Emotions in Visual Music: An Empirical Test of Two Models of Conformance and Contest for Abstract Multimedia Information Communication

A. Irene Eunyoung Lee, B. JaeSeung Jeong

A. Dong-Ah Institute of Media and Arts; B. Korea Advanced Institute of Science and Technology (KAIST)

Key words: Emotion, Perception, Multimodal, Abstract Visual Music, Multimedia Information Communication, Auditory and Visual

1. Introduction

Visual Music is inter-media through which music and visual art are truly interacted, and often they are truly united in digital media. Several previous psychological studies assert that both music and visual material bring in effects of their own when experienced in isolation. Using auditory and visual channels to present information in combination, apparent new effects get generated according to Chion (Chion, M. et al., 1994). This study investigates how emotional information in the two channels is perceived in two different conditions – the conformance and the contest condition. When the emotion information of the two media is consistent in both valence and activity (arousal) dimension is called conformance condition. In contest condition, the emotion information of auditory and visual channels is on the opposite sides of both valence and activity (arousal) dimensions. According to Somsaman (Somsaman, K., 2004), the perception of emotions follow the channel dominance model while it follows the integration model in a contest condition where perception of emotions from multimedia falls between those of the individual media.

2. Stimuli Design and Experiment

A set of stimuli (60s length in time) that consisted of four video clips, four audio excerpts, and eight multimedia combinations was created. The four video clips were designed so that two could elicit the positive emotional response and the other two could elicit the negative emotional response. The four audio excerpts were also created so that two of them to express the positive and the other two could elicit the negative emotional expressiveness. The individual media and the combination of stimuli could be separated into three conditions:

A. Individual audio and visual. There were eight stimuli – four audio excerpts and four video clips.

B. Conformance multimedia combination – Four combinations in the set were in this condition.

C. Contest multimedia combination. Four combinations in the set were in this condition.

Total of 177 students from Seoul National University, Dongah Institute of Media and Arts, and Kunyang University in Republic of Korea were recruited the effects of the effects of unified perception on audio and visual channels. We divided into four different groups of subjects to test the three conditions. Group A of 33 subjects (males: female = 10:23 with a mean age of 22.33 years) was presented to four individual audio stimuli, and then the subjects answered self-evaluation mood questionnaires to quantify the effects stimuli. In the same manner, group B of 27 subjects (males: female = 16:11 with a mean age of 24.06 years) viewed four individual visual stimuli and conducted the mood questionnaires. After the individual stimuli tests, four conformance and four contest combination stimuli were presented to two different groups of subjects to evaluate the mood questionnaires; group C ((n = 64, mean age: 22.19, sex (40:24)) evaluated the conformance multimedia combinations, and group D ((n = 53, mean age: 22.32, sex (24:29)) evaluated the contest multimedia combinations.

3. Testing Results

Questionnaires of 14 moods indices with rating scales of nine points were designed to catalog the emotional meanings of the stimuli by adopting the 13 moods described by bipolar adjectives from a previous study (Brauchli, Michel, & Zeier, 1995) with an additional index of valence (pleasant/unpleasant). We also adopted the Semantic Differential Technique (Osgood, 1957) which defines the factors of evaluation, activity, and potency to analyze the connotative meaning of percepts and to provide standard mood indices to examine the connotative meanings of our movies – four sound only contents, four video only contents, four visual–music movies of conformance type and four visual–music movies of contest type. To obtain standard mood indices, we draw the two factors of valence and arousal, we used the 'pleasant–unpleasant' factor as an assimilate valence. The 'activity' factor assimilates intensity, which is represented by scales such as tired–lively, tense–relaxed, dull–energetic, and exhausted–fresh. Finally, the 'potency' factor is a control–related index which is characterized by scales such as unsafe–safe, unbalanced–balanced, not confident–confident, and light–heavy. We had to exclude the calm–restless scale from the activity index, owing to the discrepancy in the translation of the Korean questionnaires between the pilot and subject group. The final composite indices were then re–scaled into the range [–1, 1].

3.1. Audio Only

Using the valence index and composite activity index, we obtained the ratings for each piece (mean ± STD (standard deviation)): valence: 0.70±0.32 (Audio1), 0.45±0.35 (Audio2), –0.59±0.37 (Audio3); –0.56±0.38 (Audio4); activity: 0.14±0.24 (Audio1), –0.05 ±0.31 (Audio2), 0.58±0.21 (Audio3); 0.51±0.21 (Audio4).

3.2. Visual Only

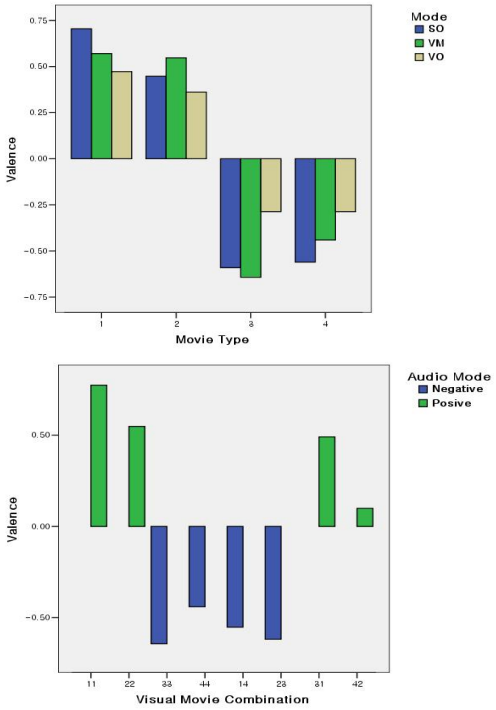
The valence and composite activity indices of visual–only, we obtained the ratings for each piece (mean ± STD (standard deviation)): valence: 0.47±0.51 (Video1), 0.36±0.49 (Video2), –0.29±0.39 (Video3); –0.29±0.51 (Video4); activity: –0.24 ±0.35 (Video1), 0.11 ±0.27 (Video2), 0.41±0.30 (Video3); 0.16±0.31 (Video4).

3.3. Conformance Visual Music

The valence and composite activity indices for conformance–type Visual Music we obtained for each piece (mean ± STD (standard deviation)) are: valence: 0.77±0.30 (VM11), 0.55±0.32 (VM22), –0.64±0.40 (VM33); –0.44±0.39 (VM44); activity: 0.01±0.20 (VM11), –0.26 ±0.21 (VM22), 0.43±0.25 (VM33); 0.55±0.26 (VM44).

3.4. Contest Visual Music

The mean analysis of valence and activity indices for each contest Visual Music movie (mean ± STD



(standard deviation)) are as shown in Fig. 1(d): valence: –0.55±0.44 (VM14), –0.62±0.48 (VM23), 0.49±0.45 (VM31); 0.10±0.45 (VM42); activity: 0.42±0.28 (VM14), 0.14 ±0.28 (VM23), 0.03±0.18 (VM31); –0.06±0.32 (VM42). .

4. Results

In general, overall tone for the emotion perception in multimedia condition seemed to be set by the auditory channel than visual channel. While it follows the channel dominance model in a contest model, it follows the integration model in a conformance condition where perception of emotions from multimedia were adjusted by the differences or enhanced by the similarity of those from the two channels.

Reference

Chion, M. (1994). *Audio-vision: sound on screen.* Columbia Univ Pr

Somsaman, K. (2004). *The Perception of Emotions in Multimedia: An Empirical Test of Three Models of Conformance and Contest.* Case Western Reserve University

Brauchli, P., Michel, C. M., & Zeier, H. (1995). Electrocortical, autonomic, and subjective responses to rhythmic audio-visual stimulation. *International Journal of Psychophysiology, 19(1),* 53-66.

Osgood, C. E., George J. Suci, Percy H. Tannenbaum. (1957). *The Measurement of Meaning:* University of Illinois Press.