Affective Responses of Visual and Auditory Stimuli for a Mathematical Theory

Hua Zhang[1,*], Wei Jiang[1,*] and Xingru Zhang[2]

[1]*Key Laboratory of Acoustic Visual Technology and Intelligent Control System, Ministry of Culture, Beijing 100024, China*
[1]*Key Laboratory of Acoustic Visual Technology and Intelligent Control System, Ministry of Culture, Beijing 100024, China*
[2]*Beijing Key Laboratory of Modern Entertainment Technology, Beijing 100024, China*
[1]*zhanghua_cuc@126.com, [1]jw@cuc.edu.cn, [2]zxr@cuc.edu.cn*

Abstract

The research tries to study humans' affective responses to superimposed costimulatory signal. Signals, including music, images, may be perceived through sound and visual stimulus. This study is interested in the perception of superimposed vibrations, whereby two or more signals are perceived simultaneously, producing a perceptual impression that is considerably different than of each signal alone, owing to the interactions between perceived stimulus vibrations that induce a coordinated percept of a mathematical theory. We assume that the correlation system is multimodal and convert them into emotional space domain. To explore between visual and auditory emotional experience sensibility relations and mapping rules, construct merging multi-disciplinary integrated intelligent emotional fusion theory system.

Keywords: Affective Responses, Superimposed Signals, Dempster-shafer evidence theory, Psychophysics

1. Introduction

In the study of emotion in music and images, there is no more meaningful than studying the relevance perception response between music and image. "The dictionary of psychology" explains: "emotion is a person to the objective things to meet their needs and the attitude of the experience". People's understanding of the objective world is from the beginning of the sense of perception. Research shows that in the visual, auditory, olfactory, taste, touch and other senses, the visual perception of human being the most important. In the external information, 80~90% is obtained through visual, auditory time. Image and music are two kinds of vivid description of the objective world, and it is the information carrier of human social activities. The human brain is like an "advanced" computer that processes the image information from the eyes and receives the sound from the ear. The theory of mirror neuron [1] confirms the expression of image, sound and other information in the human brain, which can cause the change of human emotion, but he does not reveal the relationship between the stimulating material and human emotional experience.

As is shown in the picture, the three images are the "sea" of the scene image, the object semantics and behavior semantics are basically the same, but it brings people's emotions are very different. (a) figure lets a person relax, generation of hope, (b) graph can give people the feeling of pleasure, and (c) figure gives a person is a kind of fear, deep feeling. In addition, the effects of vision and hearing on human being are not all alone in these senses. In the study of audio-visual interaction, it is found that the environmental color

will affect the auditory perception ability [2-4]. Of course, the modern art and sound film and silent film examples show this obvious impact.

Nobel prize winner HerbSimon, on the basis of the discussion of cognition, the emphasis on thinking and problem solving in general theory must contain emotional impact [5]. Emotional theorists also believe that emotion as a powerful factor in the promotion of perception, cognition, and creativity. Some conclusions of neuroscience, cognitive science, and psychology also suggest that emotions play a key role in attention, planning, reasoning, learning, memory, and decision-making. Some scientists believe that in a complex and unpredictable environment, only limited resources system will naturally lead to emotional needs, resulting in a flexible, intelligent and effective way to express multiple concerns [6-7].

The study of color vision is a branch of visual perception research. Newton's masterpiece "optics" marks the beginning of the color vision science. By Tomas Yang proposed by Helmholtz to be development of the theory of color vision, with strict logic to prove that the color is a kind of feeling. Psychology is a basic science to study the occurrence, development and the law of psychological phenomena.

Psychology refers to people's thoughts, feelings and other internal activities; is the feeling, perception, memory, thinking, emotion, will, personality, creation and so on, is the human mind to reflect the objective reality of the process.

Color psychology study proves that color is not only to people "cold", "warm" feeling, but also can give a person with "light", "heavy" feeling, "far", "near" feeling. In people's life habits, Color has formed a certain meaning of the representative. Color preferences will be influenced by different background, geographical environment, social culture and ethnic customs. People in different countries and regions of the color of the love will be a certain difference, and the same country，and the same region of different ethnic groups have different color preferences. In addition, color preferences will be affected by the background of the times, geographical environment, cultural and educational, ethnic customs, social and economic development and the impact of different social and professional groups, so it is not a fixed.

Why do some parts of the field of vision form the shape, while the rest of the field becomes the background? This is based on the results of a certain stimulus characteristic. Von Senden through the study of visual perception in patients with congenital cataract, it is proved that the distinction between the image and the background is the basic and basic link in the process of perceptual organization. At the beginning of twentieth century, the determinant factors of composition pattern are demonstrated from different angles such as proximity, continuity and symmetry by Max Wertheimer, Kurt Koffka, and Wolfgang Kohler, who were in the form of Gestalt psychology.

Sound material evoked emotions by stimulating auditory. The researchers found the impact of sound on the individual emotional experience, for example dialogue, gentle voice and gentle tone pleasant, and listening to music can be relaxed and comfortable, and sharp tone on behalf of warning, such as fire alarm siren sound. Breitenst that can produce different effects of individual emotional experience is speed and sound frequency. The study of Bradley and Lang proved that the sound can change the level of emotional arousal of the human beings, and also can induce the positive and negative emotions. Kari Kallinen found to distinguish the basic mood music evoked from the dimension of emotion, happiness and sorrow than fear, and surprise easier performance and identification, subjects of pieces of music emotion intensity and music background knowledge has no obvious relation.

In the paper, we study the single modal visual and auditory emotional experience phenomenon, to explore the emotional experience sensibility relations and mapping rules between visual and auditory, and construct merging multi-disciplinary integrated intelligent emotional fusion theory system.

(a)   (b)   (c)
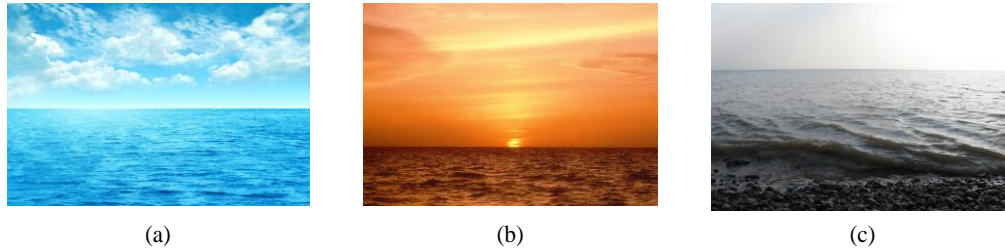
Figure 1. Different Emotion Pictures based on the Same Semantics

## 2. Responses Mechanism to Audio-visual Association

People through the perception of external stimuli, through the human brain emotional information fusion, the external stimuli to promote the change of emotional state, people in the total stimulation to achieve a new emotional state. Response model, as shown in Figure 1, S stands for the degree of external motivation, C represents the extent of the external stimulus, and P represents the basic credibility.
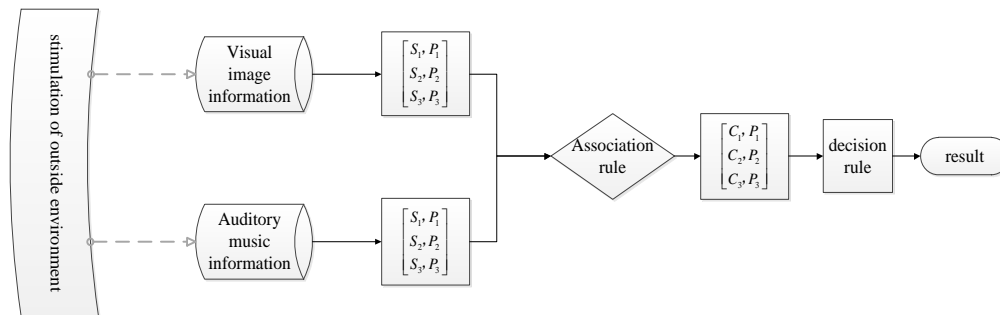
Figure 2. The Association Scheme of Audio-Visual Emotional Fusion

2.1. Affective Response Space

One of the most important problems in the construction of the affective response space is that whether or not the two or more of conflicting emotions can occur simultaneously [8]? There is still a debate about this issue in the psychological world, and this model is still considered in the emotional state of the structure. But the probability of its appearance will be small enough. Here from the simple three kinds of emotion: positive, calm, negative to construct an emotional response space, as shown in Figure 2. In this model, the origin indicates the other unknown affective states. Then, the emotional response becomes the state distribution of the emotional space. We define a three dimensional representation of the state of each affective response, and each dimension range is $(0,1)$.
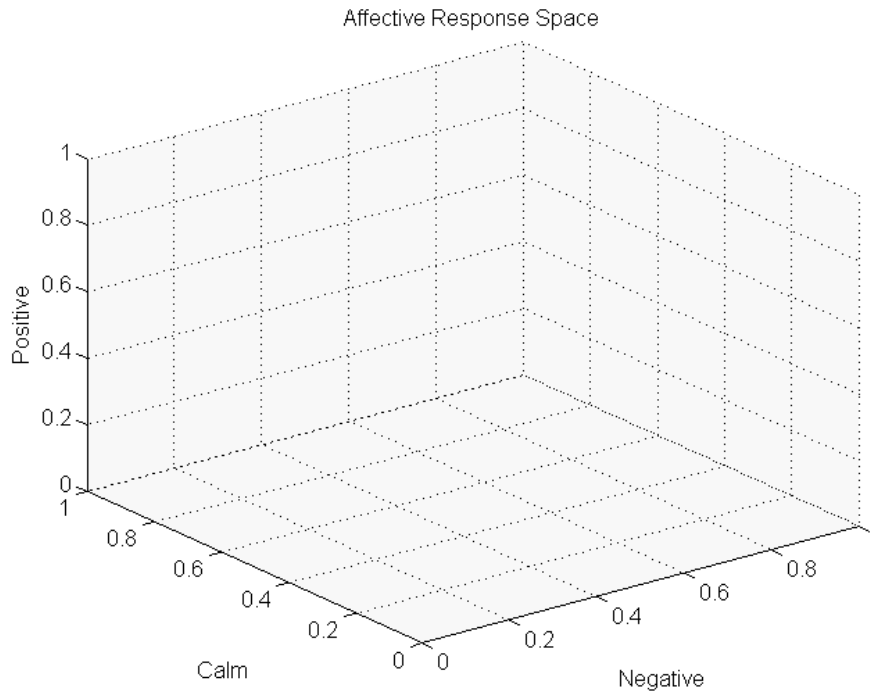
Figure 3. Affective Response Space

2.2. Dempster—shafer Evidence Theory

D-S evidence theory is proposed by Harvard mathematician A.P.Dempster when he studied the use of upper and lower probability to solve the problem of multi value mapping in 1960s [9]. Then, Then Dempster's student G.Shafer made a further development of the theory of evidence [10]. He introduced the concept of trust function, and formed a set of mathematical methods based on "evidence" and "combination" to deal with the problem of uncertainty reasoning. Dempster synthesis rules are first proposed by Dempster in the study of statistical problems. It is more intuitive and easier to obtain the prior data in the theory of evidence than the probability theory. In addition, the knowledge or data of different data sources can be integrated.

The classical evidence theory satisfies the condition that is weaker than the Bayes probability theory, that is, it does not have to satisfy the probability of additive; it has the ability to directly express "uncertainty" and "don't know". This information is expressed in the mass function, and the information is retained in the process of evidence synthesis; the theory of evidence not only allows people to give a single element of the assumption of reliability, but also can give it a subset. This is very much like the evidence gathering process of the human perception of the human brain.

(1) Basic Probability Assignment(BPA)

Let be a recognition framework, or hypothesis space. On the recognition framework, BPA is a $2^{\Theta} \rightarrow [0,1]$ m function, called the mass function, and satisfy:

$$m(\Phi) = 0$$
$$\sum_{A \subseteq \Theta} m(A) = 1 \tag{1}$$

If the proposition as elements of the framework Θ, For $\vee m(A) > 0$, A is called the focal elements of reliability function Bel.

(2) Belief Function

On the recognition framework, based on the $BPA\ m$, the belief function is defined as

$$Bel(A) = \sum_{B \subseteq A} m(B)$$

(2)

(3) Dempster's Combinational Rule

The reliability function of A is the sum of the reliability values of each subset in A. For the synthesis of reliability, $\forall A \subseteq \Theta$, m_1, m_2 is two function of mass in Θ, Dempster proposed the following synthesis rules:

$$m_1 \oplus m_2(A) = \frac{1}{K} \sum_{B \cap C = A} m_1(B) \cdot m_2(C)$$

(3)

which K is the normalized constant.

$$K = \sum_{B \cap C \neq \Phi} m_1(B) \cdot m_2(C) = 1 - \sum_{B \cap C = \Phi} m_1(B) \cdot m_2(C)$$

(4)

For $\forall A \subseteq \Theta$, a finite number of mass functions on the recognition framework $m_1, m_2, ..., m_n$, and the Dempster synthesis rules are:

$$(m_1 \oplus m_2 \oplus \cdots \oplus m_n)(A) = \frac{1}{K} \sum_{A_1 \cap A_2 \cap \cdots \cap A_n = A} m_1(A_1) \cdot m_2(A_2) \cdots m_n(A_n)$$

(5)

$$K = \sum_{A_1 \cap \cdots \cap A_n \neq \Phi} m_1(A_1) \cdot m_2(A_2) \cdots m_n(A_n)$$
$$= 1 - \sum_{A_1 \cap \cdots \cap A_n = \Phi} m_1(A_1) \cdot m_2(A_2) \cdots m_n(A_n)$$

(6)

2.3. Basic Probability Assignment based on Statistical Evidence

If a lot of evidence is based on statistical test results, it is said that this batch of evidence for statistical evidence [11]. Statistical evidence is the application of evidence theory in statistical problems, is a new processing method of Shafer for statistical problems, which is an attempt to study statistical problems with non statistical methods. The external stimuli we divided into two categories: visual (picture), auditory (Music) information. From these two aspects, we extract the emotional factors of the environment, emotional intensity S in the $[0,1]$ interval value. According to their emotional incentive factors are divided into three levels: positive incentives, moderate incentives, negative incentives. Based on statistical evidence, we determine the basic probability assignment for each music and image. The pitch is determined by the frequency, so it can be considered that the tone of the sound corresponds to the color of the light.

3. Implementation Scheme

3.1. Data Source:

**Music representative**: Chopin - b flat minor "Second Piano Sonata" third movement (Funeral March):
    1: Once we compute the spectrum;
    2: We can redistribute the energy along the different pitches;
    3: And then compute the distribution of energy along the 12 pitch classes.

Image representative: Selecting the third drawing (**c**) of Figure 1:

1: First read the image (assuming the size of the image is $N * N$ pixels) RGB values of pixels;

2: Then the 2-D Fourier transform R, G, B colors, obtaining the spectrum information of the corresponding.

Apply the statistical evidence to get basic probability assignment.
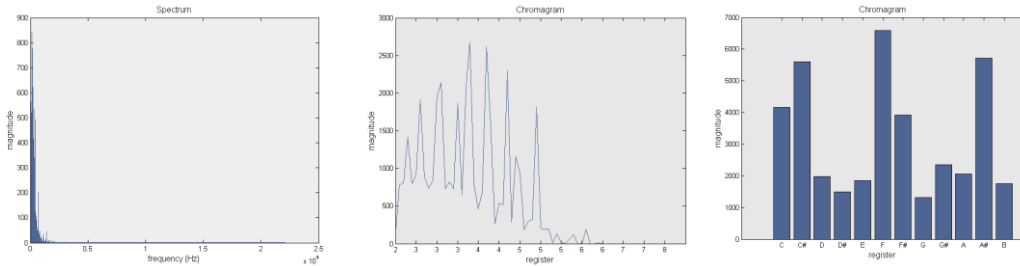
Figure 4. The Distribution of Energy along the 12 Pitch Classes
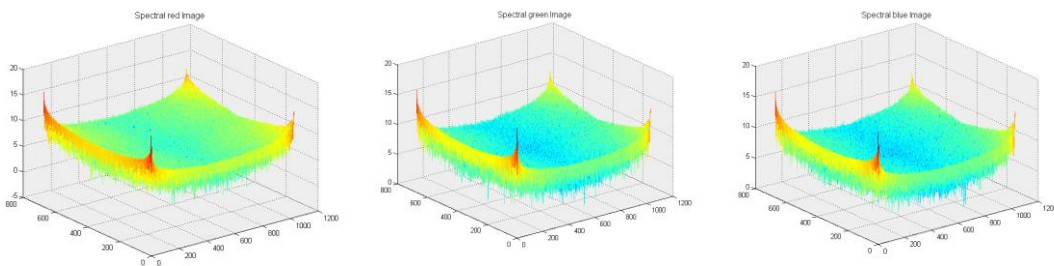
Figure 5. The Spectrum Information of the Corresponding

3.2. Mathematical Theory

Assume that S_n represents a negative incentive, S_m represents a moderate incentive, S_p represents a positive incentive, and U represents an uncertainty degree. The basic probabilities assignment determined by these three kinds of emotional information is shown in Table 1.

Table 1. Basic Probabilities Assignment Determined by these Three Kinds of Emotional Information

	S_n	S_m	S_p	U
$m_v(B)$	0.50	0.20	0.20	0.10
$m_s(C)$	0.40	0.30	0.20	0.10

Table 2. The Combination Results of B and C

	$m_v(B)$				
	$S_n(0.50)$	$S_m(0.20)$	$S_p(0.20)$	U(0.10)	
$m_s(C)$	$S_n(0.40)$	$S_n(0.20)$	$\emptyset$ (0.08)	$\emptyset$ (0.08)	$S_n(0.04)$
	$S_m(0.30)$	$\emptyset$ (0.15)	$S_m(0.06)$	$\emptyset$ (0.06)	$S_m(0.03)$
	$S_p(0.20)$	$\emptyset$ (0.10)	$\emptyset$ (0.04)	$S_p(0.04)$	$S_p(0.02)$
	U(0.10)	$S_n(0.05)$	$S_m(0.02)$	$S_p(0.02)$	U(0.01)

According to the Dempster theory, the combination results of B and C as shown in Table 2. Where $\emptyset$ represents the empty set, and the normalized constant K can be calculated as:

$$K = 1 - \sum_{B \cap C = \Phi} m_1(B) \cdot m_2(C)$$

$$= 1 - (0.08 + 0.08 + 0.15 + 0.06 + 0.1 + 0.04) = 0.49$$

(7)

Therefore, the bpaf of affective information fusion based visual (image) and auditory (music) is:

$$m_v \oplus m_s(S_n) = \frac{1}{K} \sum_{B \cap C = \{S_n\}} m_1(B) \cdot m_2(C)$$

$$= \frac{1}{0.49} \times (0.2 + 0.04 + 0.05) = 0.59$$

(8)

$$m_v \oplus m_s(S_m) = \frac{1}{K} \sum_{B \cap C = \{S_m\}} m_1(B) \cdot m_2(C)$$

$$= \frac{1}{0.49} \times (0.06 + 0.03 + 0.05) = 0.22$$

(9)

$$m_v \oplus m_s(S_p) = \frac{1}{K} \sum_{B \cap C = \{S_p\}} m_1(B) \cdot m_2(C)$$

$$= \frac{1}{0.49} \times (0.04 + 0.02 + 0.02) = 0.16$$

(10)

$$m_v \oplus m_s(\mathrm{U}) = \frac{1}{K} \sum_{B \cap C = \{\mathrm{U}\}} m_1(B) \cdot m_2(C)$$

$$= \frac{1}{0.49} \times () = 0.02$$

(11)

4. Result

4.1. Affective Responses of Visual (image 'c') and Auditory (music 'funeral march') stimuli

According to the setting of different parameters, we simulate an external stimulus in the case of audio-visual emotional superposition response process. We only consider when we receive external stimuli, the superimposed response change process in the emotional state of the individual stimuli.

In order to further simplify the problem, we set the emotional state of each dimension in the (0, 1) value. This (1,0,0) has a negative, (0,0,1) with positive, (0,1,0) with calm. For the origin (0,0,0), This is not just a calm, but also represents the other unknown emotional state. There are only three kinds of emotional time, that is, the full emotional space. For the emotional state of uncertainty, no matter what kind of emotion it will be transferred in accordance with the maximum probability of probability. According to people's experience, we assume: in a negative state while under negative incentive, the negative state will be toward the direction of negative intensity increases; in a positive state while under positive incentive, the positive state will be toward the direction of positive intensity increases; in the negative state at the same time by the opposite incentive, the emotional state is not sure, and will make people feel uncomfortable.

4.2. Affective Responses of visual (image 'c') and Auditory (music 'funeral march') Stimuli

Therefore, three music material and three image material as shown in Table 3, According to the different subjective emotions expressed in the material. Later experiments will be based on empirical information such as color and tone to analyze the status and changes of the superposition of emotional response process.

Table 3. The Material of Music and Image

|	music	image
Material 1	Zhou Haihong - Music Library - CD10-08 mildly activel -11	
Material 2	Zhou Haihong - Music Library - CD10-10 small beautiful -09	
Material 3	Chopin - b flat minor "Second Piano Sonata" third movement (Funeral March)	

4.3. Response Results Analysis

Figure 6 reflects the effective response of visual and auditory stimuli about the material 3 in Table 3.
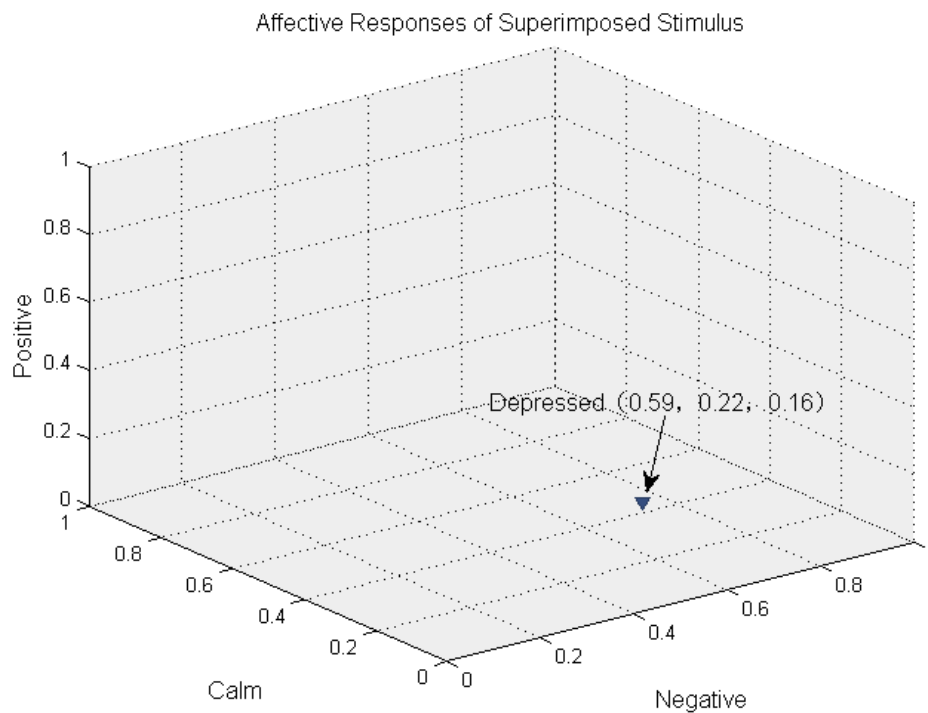
Figure 6. Affective Response (1) of Visual and Auditory Stimuli for the Mathematical Theory

(a) Response 2   (b) Response 3   (c) Response 4

(d) Response 5   (e) Response 6   (f) Cluster of Responses
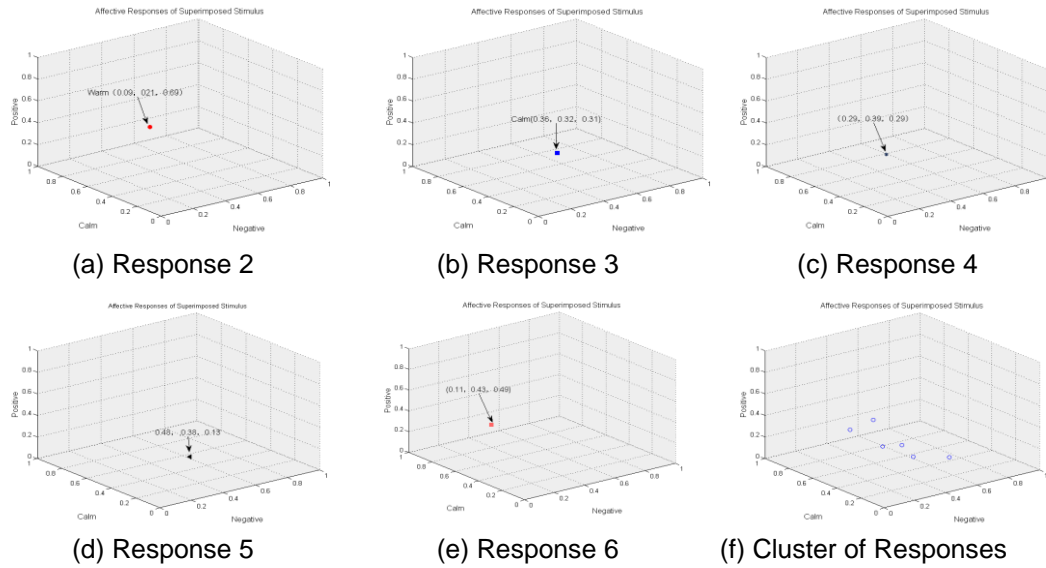
Figure 7. Affective Response of Different Superimposed Stimuli in each of the Dempster

In order to find effective changes of other stimulus superposition, we analyze the possible six results and affective spatial distribution, and the results are written in Table 4.

Table 4. The Response Results Analysis of Visual and Auditory

Affective response	Affective state	Affective trend description
response 1 Figure 6	Depressed state	The superposition response of music material 3 and picture material 3: in a negative state while under negative incentive, the negative state will be toward the direction of negative intensity increases.
response 2 Figure 7(a)	Positive state	The superposition response of music material 2 and picture material 2: in a positive state while under positive incentive, the positive state will be toward the direction of positive intensity increases.
response 3 Figure 7(b)	Calm state	The superposition response of music material 1 and picture material 1: in a calm state while under calm incentive, the calm state will be toward the direction of calm intensity increases.
response 4 Figure 7(c)	Uncertain state	The superposition response of music material 2 and picture material 3: in the negative state at the same time by the opposite incentive, the emotional state is not sure, and will make people feel uncomfortable.
response 5 Figure 7(d)	Mild negative state	The superposition response of music material 1 and picture material 3: in a calm state while under negative incentive, the calm state will be toward the direction of mild negative intensity increases.
response 6 Figure 7(e)	Mild positive state	The superposition response of music material 1 and picture material 2: in a calm state while under positive incentive, the calm state will be toward the direction of mild positive intensity increases.
cluster of responses Figure 7(f)	State distribution	It reflects the state distribution of affective response in the whole space.

The following parts mainly analyze the results of effective response, to explore the relevance of audio visual emotion based on psychology and physical feature extraction; the influence of emotion statistical characteristics on the basic probability assignment; the role of human's character in emotional response.

Analysis 1: The relevance of audio visual emotion based on Psychology.

From the angle of psychology, according to their own feelings and subjective judgments of the association, creators own emotions and experience are the main dominant factor. Synaesthesia refers to people's sense organs, namely the vision, hearing, smell, taste and touch each other through, the stimulation of a sensory organ can can cause other sensory feelings. The visual and auditory sense based, or called "audio-visual synaesthesia". This shows that the human beings have some common psychological tendency when they are dealing with the relationship between audio and video. Psychological association has a more humane and emotional characteristics, which can better reflect the people's understanding of audio-visual integration.

Analysis 2: The relevance of audio visual emotion based on Physical feature extraction.

Music in pitch, loudness, timbre, melody, rhythm and other elements can establish rich contraposition relationship with the characteristics of their visual perception of color, brightness, shape and volume. Although there is no logical association of universal and all the rules, but some of the common sense of human terms can be summed up the law. Such as the frequency of music in the color of the bright dark has an instinctive alignment between the judgments. The higher frequency sound can make people produce more brightly visual image, and the lower frequency sound can make people produce darker visual image.

Analysis 3: The influence of emotion statistical characteristics on the basic probability assignment.

Based on sensibility of D-S evidence theory, how to obtain the basic probability assignment is always closely related to the subject, as the key to the practical application, it directly affects the finally fusion accuracy and effectiveness of the decision. Therefore, the analysis of 1 and 2 is very important for the establishment of the model. In addition, the basic probability assignment method is based on statistics in the paper. To describe and deal with the uncertainty of multiple evidences in this paper, the basic idea of fuzzy set is to activate the absolute membership of the common set, which may be more beneficial to the effectiveness of the results.

Analysis 4: The role of human's character in emotional response.

Above the emotional fusion is only considered the external factors, without taking into account the internal factors of the transfer of emotion, that is, the influence of personality on the emotional response. For people with different personalities, the same external stimuli will make different emotional responses. In order to make the model more close to the true feelings of human beings. Further research will set a sensitive factor to the different degree of emotional stimuli, the matrix of these sensitive factors is the character matrix. Subsequently, the same external stimuli will produce different levels of emotional motivation.

5. Conclusion

In this paper, we propose an emotional response model to describe the change of the state of the superposition response under the double effects of visual stimuli and auditory stimuli. We use D-S evidence theory to realize the emotional fusion of visual and auditory stimuli. Introducing the psychological correlation analysis between auditory and visual, to

make the response model can be perceived by the emotional trend of superimposed stimuli. The experimental results show that the transfer of the state of the model is in line with the law of human nature. In particular, the perception of the characteristics of the material to stimulate the value of the probability assignment, making the model of the fusion result more vivid, more close to the real people's feelings.

Acknowledgment

References

[1] G. Rizzolatti and L. Craighero, "The mirror-neuron system", Annual Review Neuroscience, vol. 27, (**2004**), pp. 169-192.

[2] D. Ben-Tzur and M. Colloms, "The Effect of Maxx Bass Psychoacoustic Bass Enhanc ementon Loudspeaker Design", Preprint of Aes Convention Munic audio Eng., soc, (**2012**).

[3] BaiMR.SynthesisandImplementationofVirtualBassSystemwithaPhase-Vocoder Approach [J]. Journal of the Audio Engineering Society, vol. 54, no. 11, (**2006**), pp. 1077-1091.

[4] H. Fastl and Z. Wicker E. Psychoacoustics: Facts and Models [J]. PhysicsToday, (**1999**).

[5] H. A. Simon, "Motivational and emotional controls of cognition (1967). INMODELSOF Thought, Yale University Press, New Haven, (**1979**), pp. 29-38.

[6] R. W. Picard, Affect. Computer, MITPressLtd., United States, (**2000**).

[7] X. Hu, L. Xie, X. Liu, Z. Wang and V. Bhatnagar, "Emotion Expression of Robot with Personality", Mathematical Problems in Engineering, (**2013**).

[8] I. Hupont, S. Baldassarri and E. Cerezo, "Facial emotional classification: from a discrete perspective to a continuous emotional space", Pattern Analysis and Applications, vol. 1, (**2013**).

[9] A. Dempster, "Upper and Lower Probabilities Induced by Multivalued Mapping", Ann．Math Statist, vol. 3, no. 3, (**1967**), pp. 325- 339.

[10] G. Shafer, "A Mathematical Theory of Evidence", NJ: Princeton University Press, (**1976**).

[11] S. Rui, S. Shangyuan and G. Yunfeng, "The acquisition of basic probability assignment based on D-S evidence theory", Modern machinery, vol. 4, (**2006**), pp. 22-23.