Soundscape emotions categorization and readjustment based on music acoustical parameters

Cheng, Stone[1], Fei, Hsu[2], Cheng-Kai Hsu[3]
National Chiao Tung University
1001 University Road, Hsinchu Taiwan

ABSTRACT

This study presents an approach to analyse the inherent emotional ingredients in the polyphonic music signals, and applied to the soundscape emotion analysis. The proposed real-time music emotion trajectory tracking systems are established by machine learning techniques, music signal processing, and the integration of two-dimensional emotion plane and taxonomy as emotion recognition model. Two sets of 192 emotion-predefined training data are collected, which are popular music and western classical music respectively. Music acoustical parameters of volume, onset density, mode, dissonance, and timbre are extracted as the characteristics of music signal. Experimental results verified that different sets of training data would lead to the variation of boundaries among two emotion recognition models. This study proposed an access to environmental sound designing based on emotion recognition and psychoacoustics, especially focusing on the needs of various fields for commercial purpose or auditory atmosphere creation. The soundscape study is conducted by evaluating the effectiveness of emotion locus variation of selected urban soundscape sets blending with music signals. The simulation of playing background music in authentic field makes good use of music emotional characteristics to help people alter the emotion states and the state of mind, and further affect human behaviour and decision-making.

Keywords: Soundscape, Psychoacoustic, Emotions
I-INCE Classification of Subject Number: 79

1. INTRODUCTION

Emotional perception is the emotion expressed and conveyed by the source of the signal [1]. Discrete or basic emotional theory [2,3] focuses on persistent basic emotions, especially anger, fear, joy, disgust, sadness, and happiness. These emotions play an important role in the adaptation and adjustment of individuals to events that have a potentially important impact on their physical and psychological integrity. The antecedent of these emotions is the condition of environmental decision, which has already had a perceptual or real influence on the happiness of the individual. How people pay attention to the impact of all sound phenomena in the living environment is enhanced by the concept of “soundscape”. Handbook for Acoustic Ecology defines ‘Soundscape’ as “An environment of sound (or sonic environment) with emphasis on the way it is perceived and understood by the individual, or by a society” [4]. It thus depends on the relationship between the individual and any such environment. Daily environmental soundscapes trigger different emotions in different cultures and acoustic communities. Soundscape

[1] stonecheng@mail.nctu.edu.tw
[2] aa531078@gmail.com, [3] hsuk1011@gmail.com

emotion recognition has received increasing interest from research communities, such as perception model [5], soundscape emotion dataset [6], sound design [7], acoustic ecology [8], and urban planning [9]. The general notion that emotions generated by soundscape may have special characteristics compared with day-to-day emotions, or with emotions generated by other arts, such as music. There is a need for a model that identifies underlying dimensions of soundscape perception, and which may guide measurement and improvement of soundscape quality. Several studies are built from listening experiment [5,6]. To model and predict the perceived emotion of soundscapes, we propose music emotion data in the form of labelled audio recordings.

Flipping through the history, music has never been absent from various kinds of civilizations. Although music has evolved into different forms and genres over time, music is indeed the way that people express emotions. Expanding researches have dedicated into the field of Music Psychology and Music Emotion Recognition (MER) to explore the relationship between music and emotion [10-13]. Lu *et al*. [14] suggested the music features that evoked human emotions are "Intensity," "Timbre," "Rhythm." Lu also came up with a point that mood in the entire music piece is time-varying so the task of mood detection has extended into mood tracking. Schuller et al. [15] tried to conduct emotion recognition to the general sound events, including animals, musical instruments, nature, noisemaker, people, sports, tools and vehicles. By discussing human labelers' consistency, feature relevance and automatic regression, the correlation coefficient of "Arousal" was 0.49, and the correlation coefficient of "Valence" was 0.61, which indicated the feasibility of analyzing emotion ingredients of general sound events with AV axes.

Emotion model can be divided into two approaches, including categorical approaches and dimensional approaches. Categorical emotion model was proposed on the basis of "Basic Emotions", which are consisted of few emotions that are inherent in human, other emotions can be derived from basic emotions. Distinct from categorical approaches which classified emotion responses into category, dimensional approaches regarded emotion state as a transition process, among which the two dimensional emotion plane is the most popular. Figure 1 shows emotion models of various dimensional approach researches and with basic emotions attached. This study has resampled the two approaches of emotion model and designated "Energy", "Stress" as the two axes of emotion plane. The energy axis covered the concept of "Arousal", "Activation" and "Activity", whereas the stress axis merged the concept of "Valence", "Tension" and "Displeasure". With four mood adjectives attached on the emotion plane, the emotion recognition model adopted in this research is shown as Figure 2.

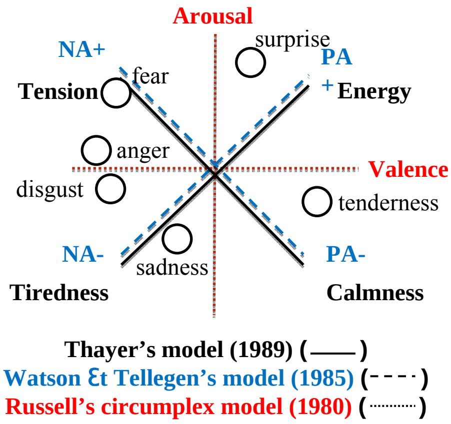


Figure 1. Emotion Model of various researches

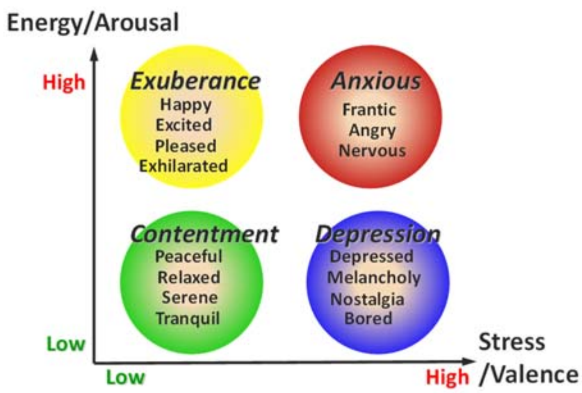


Figure 2. Emotion Model resampled in this work

This study aimed at quantifying and visualizing the emotion response evoked by music pieces and soundscape sets. By integrating the theory of psychology, machine learning techniques, and music database, a real-time music mood transition tracking system is established and applied to the soundscape emotion analysis. This paper is an attempt to design appropriate auditory atmosphere of selected soundscape based on psychoacoustic.

2. SYSTEM ARCHITECTURE

Feature extraction from music signal is one of the most popular technique used in content-based music information retrieval to figure out the characteristics of the music pieces. A variety of studies in the area of Music Psychology have discussed about the relationship between music features and AV axes, which is listed in Table 1. Figure 3 illustrates the block diagram of the proposed system, the blue path is the training mode and the red path is the testing mode. Two sets of training data are consisted of 192 classical music clips and 192 popular music clips with human labeled emotion tag in four emotion groups, “Pleasant”, “Solemn”, “Agitated” and “Exuberant” respectively. The length of each music clips is 30 seconds and the format is WAVE file, sampling frequency 44.1 kHz, Stereo, 16 bits resolution. The input music data would be down sampled and framed in the step of preprocessing in favor of the computation afterward. Five music features are extracted in this research to progressively characterize music-evoked emotions, including onset density, dissonance, timbre, mode and volume.

The function of “Emotion score counting” is used to simulate the process of how human emotions are being evoked while music listening. The concept is that the listeners’ mood is triggered by the feelings of previous time period, hence we separate the process into two steps, present emotion score and cumulative score counting. As shown in Equation 1, p_t is a two dimensional function which is specified as present emotion score, x axis and y axis stand for “Stress” and “Energy” respectively, f is the index of five different music features extracted in this study，$S_f(t)$is the strength of f-th feature at the time t, $wx(f)$, $w_y(f)$ are the weights of each feature on the x axis and y axis.

$$p_t(x,y)=\sum_{f=1}^{5}\left\{\left[w_x(f)S_f(t)\bar{x}\right]+\left[w_y(f)S_f(t)\bar{y}\right]\right\} \quad (1)$$

Table 1 - *Music Features related to the Arousal and Valence Emotions*

		Onset Density	Volume	Timbre	Mode	Dissonance
Juslin *et al.* [10]	Arousal	✓	✓	✓		
	Valence	—	—	✓		
Laurier *et al.* [16]	Arousal	✓	✓		—	✓
	Valence	—	—		✓	—
Schubert [17]	Arousal	✓	✓	✓	—	✓
	Valence	—	—	✓	✓	✓
Combined in this research	Arousal	✓	✓	✓	—	✓
	Valence	—	—	✓	✓	✓

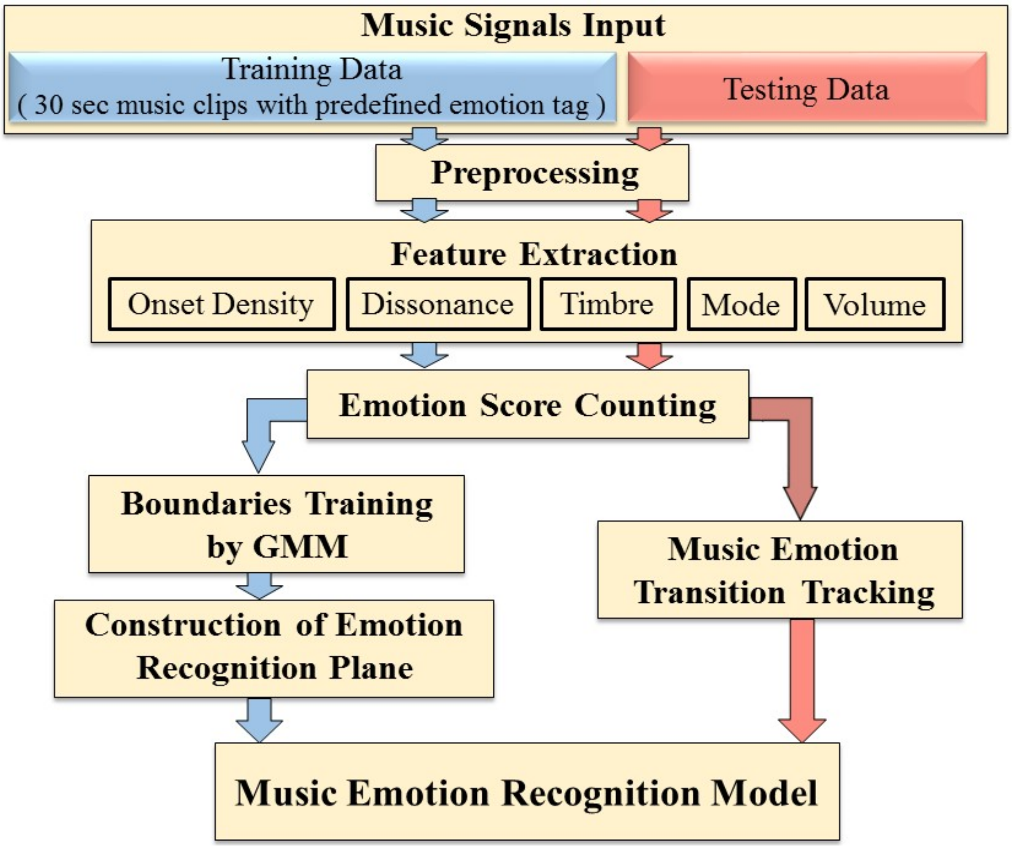


Figure 3. *System flow chart*

As shown in Figure 4, cumulative score counting process p_t multiplies present emotion scores with decay function $d(t)$, representing the release of influence of former time period on the emotion perception.

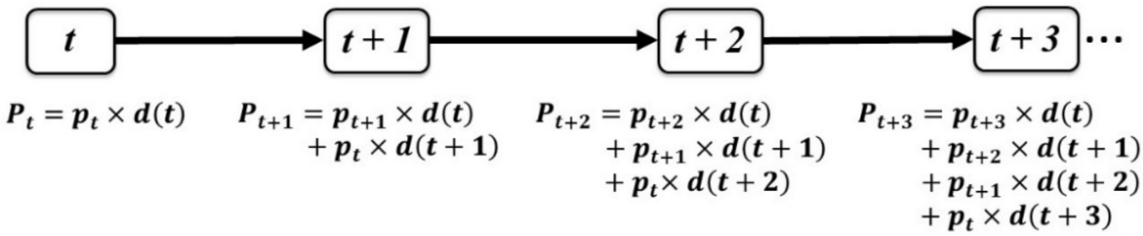


Figure 4. *Cumulative Score Counting*

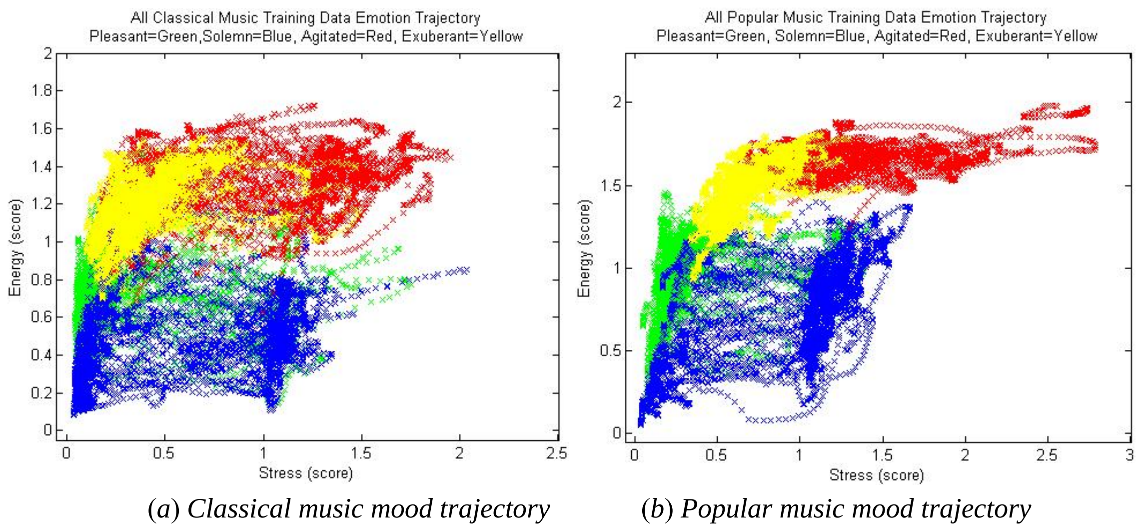
3. CONSTRUCTION OF EMOTION RECOGNITION MODEL

Music emotion scoring method is used to simulate the process of music listening. Its concept is that the listeners' emotion is influenced by the feeling at the previous time period, and as the time goes on, the emotional feeling will also release gradually. According to the scoring process, we can mark the final emotion scores of training data on the "Energy" and "Stress" formed emotion plane showing as the trajectory of each music clip. Either for classical music training data, or for popular music training data, the music features extracted in this research perform better on distinguishing different emotion-tagged music groups on energy axis than on stress axis. The result is not surprising since previous researches of MER have indicated that arousal is strongly linked with features such as volume and tempo, leading to the objective assessment on the relationship between music and "Energy". Lacking of solid connection between music features and valence, leaving the assessment of

“Stress” become subjective and hard to quantify [17].

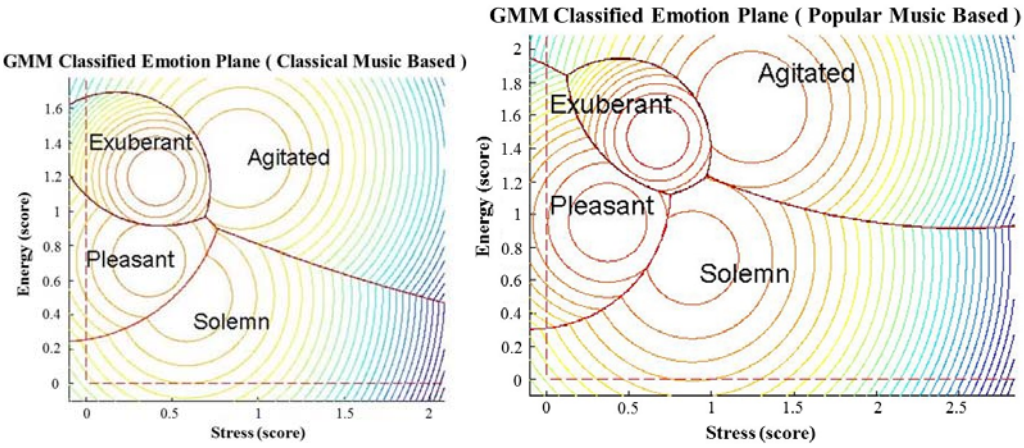
The locus of popular music training data is more distinguishable on the “Stress” axis, however, most of the locus of classical music training data is overlapped. The reason is that the music features attributed to the “Stress” axis include timbre, mode, and dissonance, abundant of electric instruments and synthesized sound are usually employed in popular music. The timbre of electric guitar can be varied with the aid of effect such as flanging, phasing and modulation, and may distort the sound signal to cause the dissonance. The negative emotion can be conveyed out easily through this sound distortion process. The timbre and the harmony of classical music are more monotonous compared to popular music, resulting in the difficulty of shaping contrast among different emotion states’ music.

Gaussian Mixture Model (GMM) algorithm has capability of accurately present data distribution by adding various Gaussian function with different weights. According to the trajectory distribution showed in Figure 5, GMM computes the Probability Density Function (PDF) of each mood locus, finally the optimal boundaries between four emotion states are carried out and the emotion planes are shown in Figure 6. Figure 5(a) shows that the trajectories of “Pleasant” and “Solemn” from “classical” music training data have more overlap area than the trajectories of “Pleasant” and “Solemn” from “popular” music training data, which results in the high similarity of PDF as shown in Figure 6(a). Figure 7 illustrates an example mood tracking result of a testing music piece, “1812 Overture” composed by Tchaikovsky. In the figure, the emotion ingredients in the piece is consisted of 19% of Pleasant, 72% of Solemn, 7% of Agitated and 2% of Exuberant with pop song based model while in the classical music based model is 22% of Pleasant, 56% of Solemn, 19% of Agitated and 3% of Exuberant. This study utilized the model trained by popular music training data for analyzing soundscape emotions since the trajectories of popular training data have less overlap for four emotion states, which is beneficial on emotion recognition. Also, the sound effects used in popular music are wider than in classical music, which has narrow frequency components with general sound events.



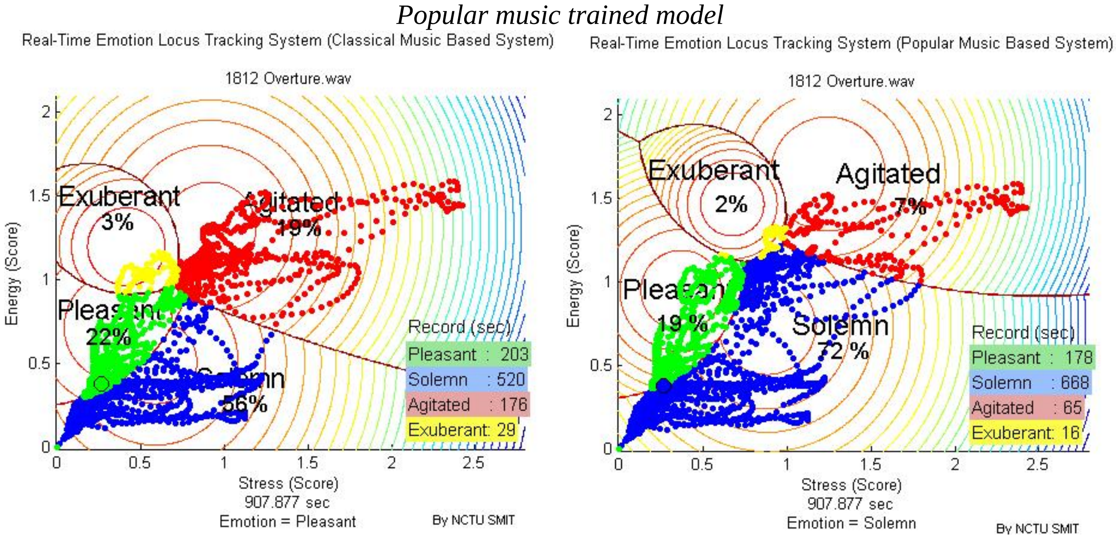
(*a*) *Classical music mood trajectory* (*b*) *Popular music mood trajectory*

Figure 5. Training data emotion trajectory distribution. (*a*) *Classical music training data* (*b*) *Popular music training data*



(*a*) (*b*)

Figure 6. GMM demarcated emotion plane and margins (a) Classical music trained model (b) Popular music trained model



(*a*) (*b*)

Figure 7. Illustration of emotion trajectory tracking results of a testing music piece, "1812 Overture" composed by Tchaikovsky. (a) Tracking in classical music based system (b) Tracking in popular music based system

4. SOUNDSCAPE EMOTION READJUSTMENT EXPERIMENTS

Soundscape is a study of the interaction between sound and human and space environment. This study combines the results of emotional analysis with the needs of specific field auditory environment design from a psychological perspective. Explore the impact of environmental soundscapes in everyday life on human emotional state.

In this experiment, several sound field fields were selected, and the field recording was performed by hand-held recorder. The volume intensity of the sound scene audio file was normalized, and the emotional mood tracking system was analyzed by the music emotion trajectory tracking system established in this study. Then, the auditory environment of individual soundscape fields is created to study the changes of emotional trajectory after mixing with specific music. In the real listening environment, the human emotion reaction is changed by playing music. The research process is shown in Figure 8. Future potential market applications include the mood of the shopping mall environment to create a desire to increase consumers' willingness to purchase, to change the mood of the workplace, to improve employee productivity, and to change the emotional mood of medical institutions to appease the industrial and commercial use of patients with anxiety and anxiety.

In this study, the selected restaurants and store soundscapes were analyzed. Various shopping malls, department stores, chain fast food restaurants, and theme restaurants have sprung up, and the number of people shopping and dining has increased. How to make the consumption environment more in line with its purpose through music has become a topic of research. Many studies have confirmed that emotional behavior can be influenced by music and thus affect consumer behavior. In exploring the influence of background music played by the restaurant on the consumption behavior of customers to the store, it is pointed out that slow-paced music can prolong the time that customers stay in the restaurant and increase the consumption expenditure. At the same time, the gentler and more beautiful music also provides a more comfortable dining environment [18][19][20]. Taking the student restaurant on the university campus as the research object, it was found that whether playing classical or popular music would increase students' desire for consumption, and the turnover did increase [21].

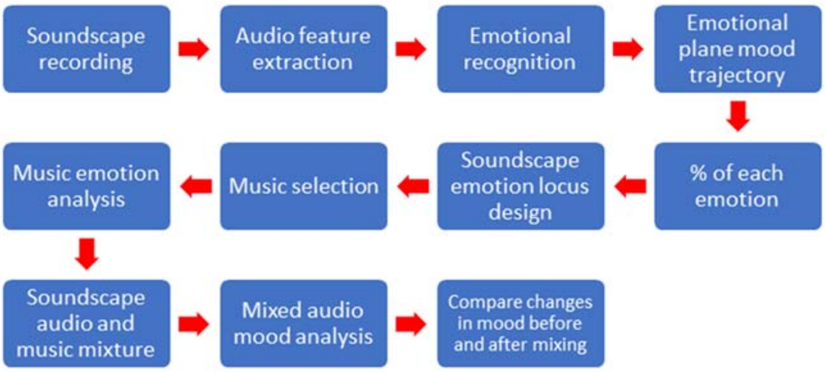


Figure 8. *Soundscape emotion analysis process*

For the retailers, supermarkets and other places where people purchase daily necessities, the relevance of music and consumer behavior, many research results show that the impact of playing music in the store on the turnover is positive[22-27]. Especially the slower pace of the concert allows the customer to stay longer. In addition, playing a less familiar track of a customer will also lengthen the time it takes to stop because humans tend to focus on new, unknown things, when they hear a piece of music that they have never heard before. Customers will increase the time spent in the store because they might want to listen to the whole song or the entire album. Related research also shows that most people's moods are pleasant when shopping, but sometimes they have anxiety and anxiety. It is more appropriate to play soothing music. On the basis of the music emotion trajectory tracking system, we selected mealtime restaurant soundscape as testing data. Figure 9 shows the emotion trajectories of testing data, the whole locus are in “Solemn”, which indicates that the emotion responses triggered by restaurant soundscape are stressful and negative. Figure 10 shows the tracking result of “*Beat It*”, the mood locus are mostly in the range of “Exuberant”. By blending the restaurant soundscape with “*Beat It*” to simulate the realistic auditory atmosphere in restaurant with music being played, as shown in Figure 11, the emotion ingredients of “Solemn” have descended from 100% to 53%, while 37% of trajectories lie in “Exuberant” and 11% of trajectories are in the range of “Pleasant”, which makes the atmosphere of restaurant more comfortable and eased.

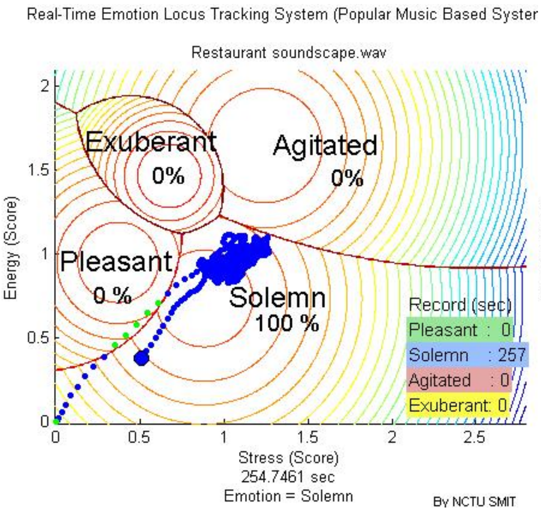


Figure 9. *Emotion trajectory of selected soundscape：Restaurant*

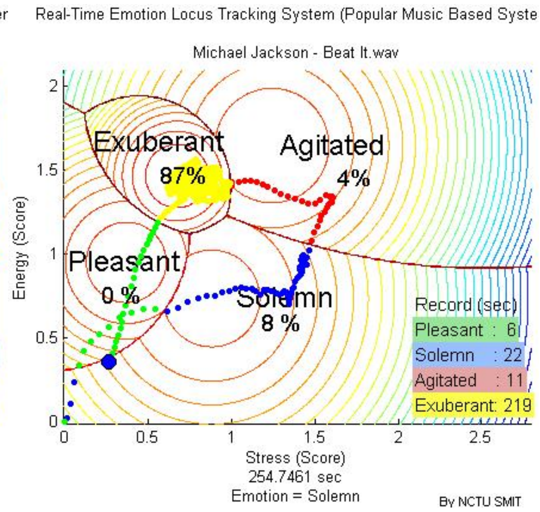


Figure 10. *Emotion trajectory of music：“Beat It”by Michael Jackson*

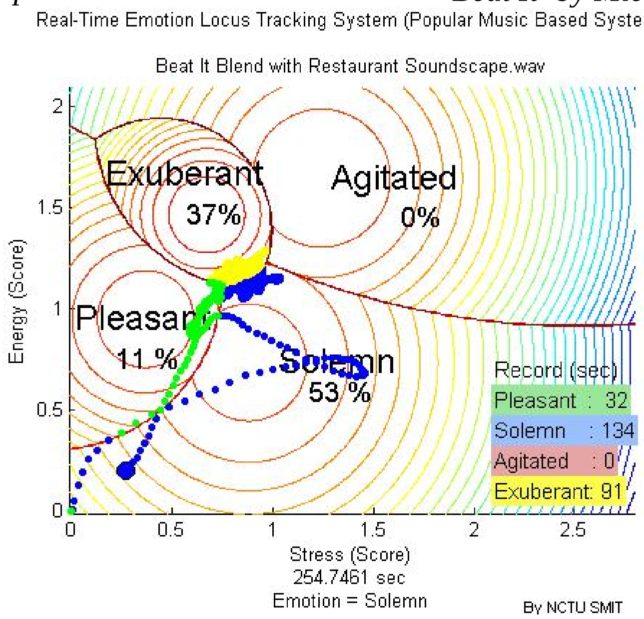


Figure 11. *Analysis on ratings of emotional responses to Restaurant soundscape blended with “Beat It” music by Michael Jackson.*

Soundscapes vary over time (minutes/hours) and this can affect people’s perception of them. The soundscape survey procedure is carried out by conducting a soundscape emotions evaluation and alteration by a blend of music signals to compare the results from proposed system analysis the actual listener experience. In testing mode, the proposed approach is evaluated with our testing database. In this framework it is possible to design for particular and predictable psychological effects by tracking the emotions of soundscape. The annoyance sound may be attributable to the sonic environment to understanding soundscape quality, allows one to engage in soundscape design for quality of life. We first present the performance of soundscape emotion tracking locus on selected soundscape sets. Figure 12 shows the proceeding of test soundscape drawn the locus of arousal emotions. The test soundscape is rivulet sound with cicadas chirping. The locus of soundscape’s emotion are in “Anxious” mostly (56 secs), which can be connected to results regarding noise sensitivity. Distal situational awareness is predominantly determined by the loudest (foreground) sound events and proximal situational awareness by the subtle (background) sounds [12]. We choose a music clip with its emotion locus is in the range of “Contentment”, as shown in Figure 13. The emotion locus of blended audio signals is shown in Figure 14. Our research indicates that an annoying sound is the main reason of its annoyance emotion. In particular people complained because of its constant, frequent, or unpredictable presence or because it had particular source properties. The blended soundscape experiment suggests that more subtle interventions that address the attention attracting properties might often be possible. For example, it may make singing sounds more audible and in doing so shift the appraisal of the sonic environment towards the “Contentment” emotion. The loudness effect of this intervention may be small, while its well-being effect may be important.

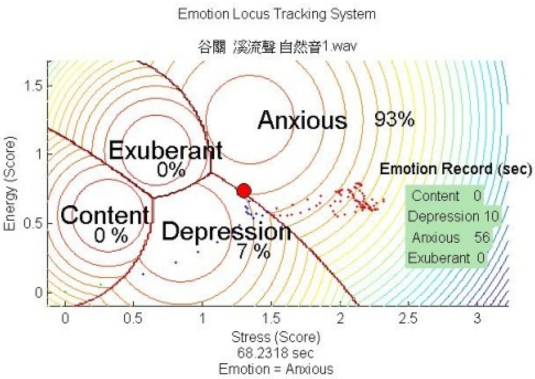


Figure 12. *Emotion locus of soundscape sample: Rivulet Stream-Nature Sounds.*

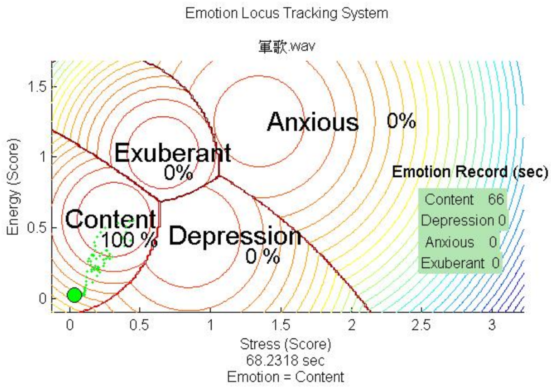


Figure 13. *Emotion locus of singing song sample*

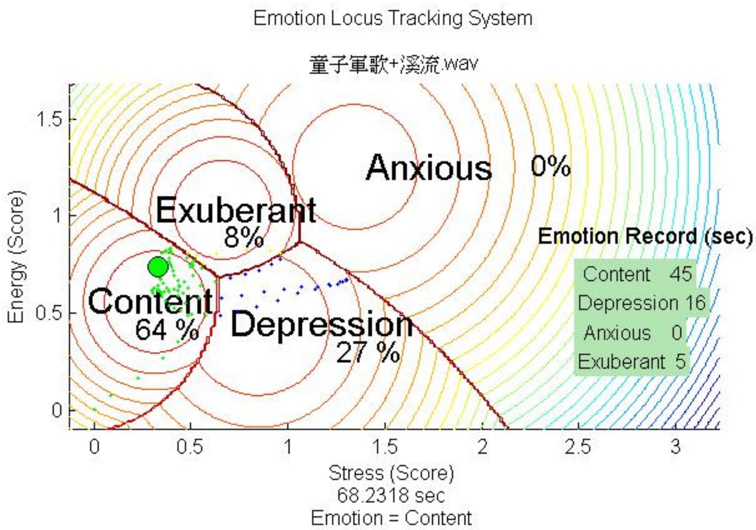
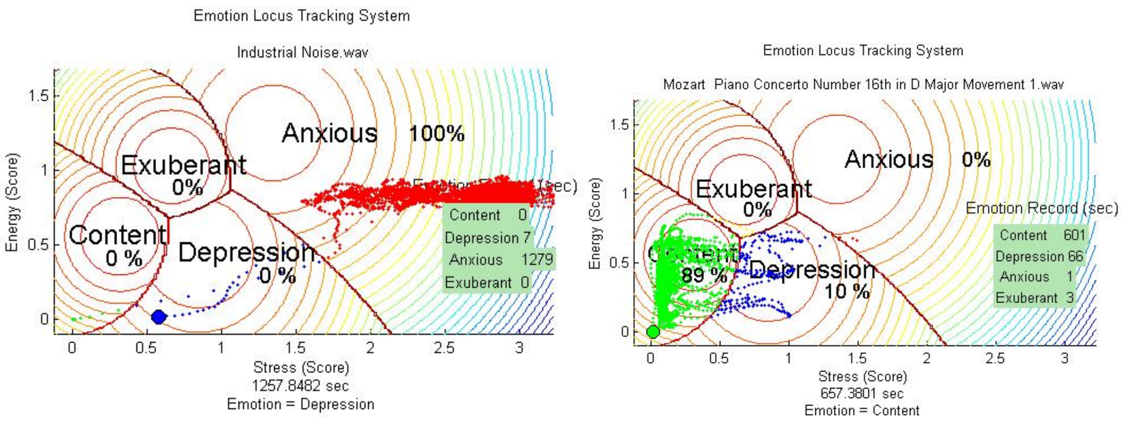


Figure 13. *Analysis on ratings of emotional responses to Rivulet Stream-Nature Sounds blended with singing song.*

Many workers on the factory found themselves in a noisy environment. The soundscapes ambience indeed affect peoples’ emotion response. This research aims at studying the interaction between sound, environment and people, especially focuses on figuring out how soundscapes influence human’s emotion state. Figure 22(*a*) shows real-time soundscape emotion for factory noise and Mozart Piano Concerto Num.16 in D Major in Figure22(*b*). The emotion of factory noise indicates that an annoying sound is the main reason of its annoyance emotion (100% 1279 sec). In particular people complained because of its constant, frequent, or unpredictable presence or because it had particular source properties. A *Mozart Piano Concerto Num.16 in D Major* music with its emotion locus (89% Content) shows in Figure 22(*b*) is mixed into factory noise soundscape. Figure 23 shows the result of soundscape for factory noise blended with Mozart Music. 100% of anxious emotion in factory noise was descended to 1%.



(*a*) (*b*)

Figure 22. Real-time Soundscape Emotion Evaluation for Factory Noise (a) and Mozart Piano Concerto Num.16 in D Major (b).

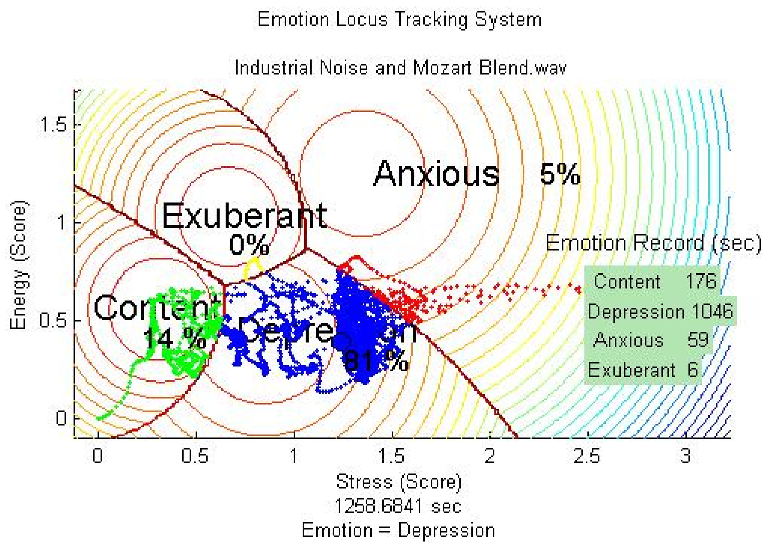


Figure 23. Real-time Soundscape Emotion Evaluation for Factory Noise Blended with Mozart Music

5. CONCLUSIONS

This study proposes a method for analysing the emotional components caused by music signals and applies them to the analysis of soundscape emotions. The proposed system integrates multiple emotion models into a two-dimensional emotional space. By analysing the five audio features of 192 music clips, an emotional classification model is established to construct a two-dimensional analysis of the emotional state. Each feature is analysed using frames of different lengths. The system uses the Gaussian Mixture Model (GMM) to classify the four emotions in the emotional plane and plot the variation of the emotional components caused by the music signal. It is used to divide the boundaries of the four emotions of "excitement", "depression", "satisfaction" and "anxiety" on the emotional plane, and establish a graphical interface of the emotional trajectory on the two-dimensional emotional model, indicating the tracking of emotional transition. In this study, the emotional trajectory analysis of the soundscape is performed on the selected soundscape set, and the appropriate music is selected by the soundscape scene to perform the change of the emotional trajectory of the soundscape to evaluate the effectiveness of changing the mood by the mixed music signal. Preliminary evaluations show that the proposed algorithm yields good results.

6. REFERENCES

1. K. Kallinen, N. Ravaja, "*Emotion Perceived and Emotion Felt: Same and Different*" , Musicae Scientiae, vol. 5, no. 1, pp. 123-147, 2006.
2. Ekman, P., "*An argument for basic emotions. Cognition & Emotion*", 6, 169–200 (1992).
3. Izard, C. E., "*Basic emotions, natural kinds, emotion schemas, and a new paradigm*", Perspectives on Psychological Science (2007), 2, 260–280.
4. Truax, B., "*Handbook for Acoustic Ecology*", 2nd ed. 1999 Cambridge Street, Vancouver, Canada; available from http://www.sfu.ca/sonicstudio/handbook/.
5. Östen Axelsson, Mats E. Nilsson, and Birgitta Berglund, "*A principal components model of soundscape perception*", J. Acoust. Soc. Am. 128,5, November 2010.
6. Jianyu Fan, Miles Thorogood, Philippe Pasquier, "Emo-Soundscapes: A Dataset for Soundscape Emotion Recognition", 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII) pp.196-201.
7. M. Thorogood and P. Pasquier, "Computationally Generated Soundscapes with Audio Metaphor," in International Conference on Computational Creativity, 2013, pp 1-7.
8. W. Ma and W. F. Thompson, "Human Emotions Track Changes in the Acoustic Environment," PNAS, vol.112, no. 47, pp. 14563-14568, 2015.
9. B. Davies, et al. "A Positive Soundscape Evaluation Tool," in Proceedings of the 8th European Conference on Noise Control-Abstracts. Institute of Acoustics, 2009.
10. Patrik N. Juslin and John A. Sloboda (Eds.), "*Handbook of Music and Emotion: Theory, Research, Applications*", Oxford University Press, 2010.
11. K. Hevner, "*Experimental studies of the elements of expression in music*," The American Journal of Psychology, Apr. 1936; 48(2):246-268,.
12. Robert E. Thayer, "*The Biopsychology of Mood and Arousal*", Oxford University Press, 1989.
13. M. Zentner, et al., "*Emotions evoked by the sound of music: characterization, classification and measurement*" , Emotion, Aug. 2008;8(4):494-521,.
14. L.Lu, D.Liu and H.-J. Zhang, "*Automatic mood detection and tracking of music audio signals*", IEEE Transactions on Audio, Speech, and Language Processing, Jan. 2006;14(1):5-18,.
15. B. Schuller, S. Hantke, F. Weninger, W. Han, Z. Zhang and S. Narayanan, "*Automatic recognition of emotion evoked by general sound events*" , ICASSP 2012; 25-30 March 2012;Kyoto, Japan 2012. pp.341-4.
16. Cyril Laurier, Olivier Lartillot, Tuomas Eerola and Petri Toiviainen, "*Exploring relationships between audio features and emotion in music*", Proc ESCOM 2009; 12-16 August 2009; Jyväskylä, Finland 2009.
17. E. Schubert, "*Measurement and Time Series Analysis of Emotion in Music*", University of New South Wales, Ph.D Dissertation, 1999.
18. G. C. Bruner II, "*Music, Mood, and Marketing*", Journal of Marketing, Vol. 54, No. 4, pp. 94-104, October 1990.
19. R. E. Milliman, "*The Influence of Background Music on the Behavior of Restaurant Patrons*", Journal of Consumer Research, Vol. 13, No. 2, pp. 286-289, September 1986.

20. C. Caldwell, S. A. Hibbert, "*The Influence of Music Tempo and Musical Preference on Restaurant Patrons' Behavior*", Psychology & Marketing, Vol. 19, No. 11, pp. 895-917, November 2002.
21. A. C. North, D. J. Hargreaves, "*The Effect of Music on Atmosphere and Purchase Intentions in a Cafeteria*", Journal of Applied Social Psychology, Vol. 28, Issue. 24, pp. 2254-2273, December 1998.
22. R. E. Milliman, "*Using Background Music to Affect the Behavior of Supermarket Shoppers*", Journal of Marketing, Vol. 46, No. 3, pp. 86-91, 1982.
23. R. Yalch, E. Spangenberg, "*Effects of Store Music on Shopping Behavior*", Journal of Consumer Marketing, Vol. 7, Iss.2, pp. 55-63, 1990.
24. R. Yalch, E. Spangenberg, "*The Effects of Music in a Retail Setting on Real and Perceived Shopping Times*", Journal of Business Research, Vol. 49, Iss.2, pp. 139-147, August 2000.
25. L. Dubé, S. Morin, "*Background music pleasure and store evaluation: Intensity effects and psychological mechanisms*", Journal of Business Research, Vol. 54, Iss.2, pp. 107-113, November 2001.
26. E. R. Spangenberg, et al., "*It's beginning to smell (and sound) a lot like Christmas: the interactive effects of ambient scent and music in a retail setting*", Journal of Business Research, Vol.58, Iss.11, pp. 1583-1589, November 2005.
27. M. Morrison, "*In-store music and aroma influences on shopper behavior and satisfaction*", Journal of Business Research, Vol. 64, Iss.6, pp. 558-564, June 2011.