Temporal features extraction for the binaural soundscape samples

Daiwei WANG[1]; Zhiyong DENG[2]; Xinxin LI[3]; Aili LIU[4]

[1, 2, 3] College of Music, Capital Normal University, China

[4] College of Resource Environment and Tourism, Capital Normal University, China

ABSTRACT

Soundscape, earlier proposed by the Canadian composer and ecologist R. Murray Schafer in 1960s, contains a large amount of environmental information and event information for the expression of specific historic areas, the survival status of a particular nation or people, and especially, the native environment has important significance. 22 binaural recorded soundscape samples collected from Guangxi Zhuang Autonomous Region of China and their temporal features, included IACF, IACC, temporal envelope, acoustical dynamic were introduced in this paper. The information of native location, sound event and the above temporal features are also highly corresponded to the keynote, sound signal and soundmark of the soundscape. The preliminary results in this paper can provider an important theoretical and practical significance for the further extraction and analysis for the acoustical parameters of soundscapes.

Keywords: Soundscape, Binaural, Temporal feature   I-INCE Classification of Subjects Number: 56.3

1. INTRODUCTION

Soundscape, earlier proposed by Canadian composer and ecological scientist R. Murray Schafer in 1960s, is a sound environment emphasized on the individual or social perception and understanding. Soundscape have three main factors: keynote, sound signal and soundmark (1). Keynote sounds expressive, and most of keynotes are natural sounds such as wind, water, sound, animals chirping and traffic noise. Sound signals can cause noticed sound events, such as artificial sounds, bell tower and whistle sound of cars. Soundmark sounds that full of characteristics, it can be not only the natural sounds but also artificial sounds. On one hand, soundscape is cultivating people to listen carefully to the surrounding sounds, and on the other hand, soundscape design and development could improve the surrounding sound environment based on the investigation, analysis and evaluation of the original sound. Meanwhile, as a interdisciplinary subject, soundscape research is involved in music, acoustics, psychology and sociology. Recently, many of papers were focused on the historical and cultural significance of soundscape, soundscape anthropology, soundscape composition, soundscape in tourism geography and other related fields (2, 3, 4, 5, 6). Those forward-looking and enlightening researches show that the essence of soundscape is original expectations from the local people, the native ecological environment, society, history and culture.

Due to the previous studies, based on 22 original binaural recorded soundscape samples, the extraction algorithms of four temporal features, included RMS temporal envelope, SPL dynamic based on the sound pressure level difference, interaural cross-correlation function (IACF) and interaural cross-correlation coefficient (IACC) were introduced in this paper. At the same time, an rough analysis for the relationship between these four temporal features and the three main factors of soundscapes also was involved at the end of this paper.

---

[1] davidyx@qq.com

[2] dengzhiyong@gmail.com

[3] 571984399@qq.com

[4] beyondtour@163.com

* Supported by the Youth Funding Project of Humanities and Social Sciences Research Foundation from the Education Ministry of CHINA, named *Case Studies for the soundscape ecology and cultural Heritage of Zhuang villages* (No. 11YJCZH026), and the Research Funding Project for Talents Training Plan of Beijing, named *The Anthropology of Soundscape, a sound history of the urban and rural areas* (No. 012135407704).

2. RESEARCH OBJECT AND SOUNDSCAPE SAMPLES ANALYSIS

In this paper, 22 binaural recorded soundscape samples collected from six different villages or areas in Guangxi Zhuang Autonomous Region of China. Sample 1-7, 11-15 were collected in Madu village of Baise City, sample 8-10 were collected in Longji village of Guilin City, sample 16,17,19 were collected in Yu-ma Park of Liuzhou City, sample18 was collected in Guangming Road of Liuzhou City, sample 20,21 were collected in Dongmen Park of Liuzhou City, and sample 22 was collected in Liuhou Park of Liuzhou City. All samples were collected on the native locations (as figure 1) without any artificial distortions, and all the information (as table 1) from these samples are adequate to reflect the acoustic characteristics of the native environments.

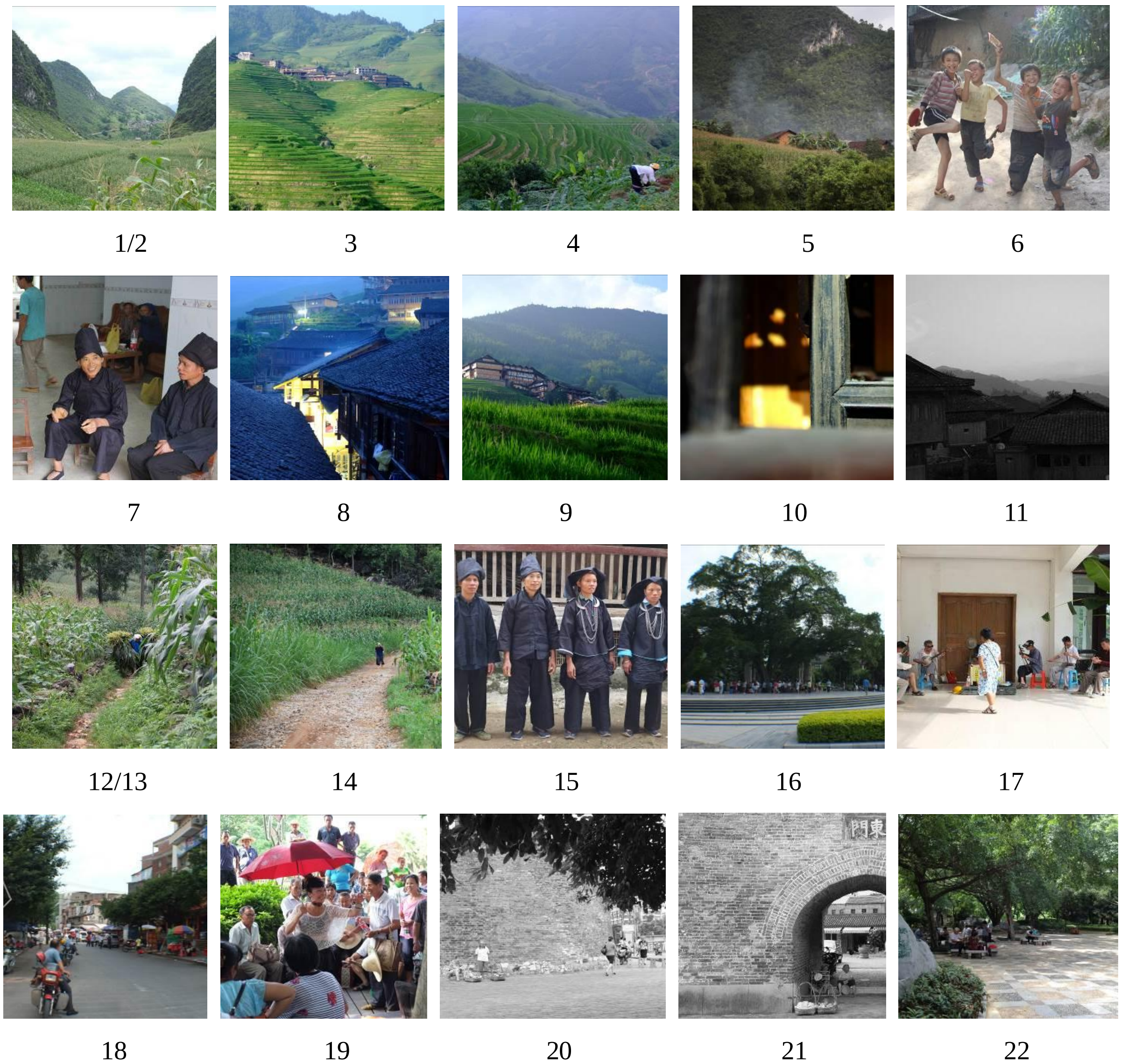


Figure1 – The native locations of the 22 soundscape samples

Table 1 – Description of the soundscape samples

No.	Duration	L_{eq}(dBA)	Description
1	00:13	48	Madu village, Baise, morning, sound of insects
2	03:49	60	Madu village, Baise, morning, from residence to entrance of the village, crowing.

3	00:54	76	Madu village, Baise, morning, crowing and birdsong, field work.
4	00:38	55	Madu village, Baise, morning, sound of file work and sounds of conversation
5	00:47	40	Madu village, Baise, morning, quiet.
6	05:50	72	Madu village, Baise, dusk, children's laughter.
7	06:44	70	Madu village, Baise, interviews.
8	05:34	57	Longji village, Guilin, environment sound.
9	07:50	70	Longji village, Guilin, environment sound.
10	01:37	50	Longji village, Guilin, wooden house, sound of footsteps on wood floors.
11	05:16	55	Madu village, Baise, evening, environment.
12	00:48	69	Madu village, Baise, morning, beside the field, course by local people.
13	05:16	75	Madu village, Baise, morning, talk to local people, birdsong and environment sound.
14	01:36	75	Madu village, Baise, morning, talk with a woman in field.
15	08:00	79	Madu village, Baise, room, 2 men and 2 women, antiphonal singing Zhuang.
16	06:22	65	Yu-ma Park, Liuzhou, tango.
17	04:43	85	Yu-ma Park, Liuzhou, small national orchestra with vocal.
18	07:33	73	Guangming Rd, Liuzhou, north to south, street vendor, traffic noise, music.
19	08:57	75	Yu-ma Park, Liuzhou, three antiphonal singing.
20	02:30	74	Dongmen Park, Liuzhou, street vendor, traffic noise, music, environment sound.
21	01:37	69	Dongmen Park, Liuzhou, roadside, the elderly people have a rest and singing.
22	04:32	65	Liuhou Park, Liuzhou, the elderly people play chess and talk to each other, traffic noise.

3. TEMPORAL FEATURES EXTRACTION ALGORITHMS OF SOUNDSCAPE

3.1 Peak extraction algorithm and RMS temporal envelope

In order to extract the RMS temporal envelope and the SPL dynamic, we use the following peak detection algorithm to extract the peaks of each sample. Firstly, the sound signals were cut by a fixed-width rectangular window without changing the amplitude of the signal. Then the peak of each windowed segment was extracted as figure 2.

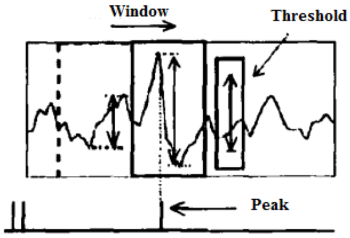


Figure 2 - schematic diagram of peak Extraction algorithms (7)

According to figure 2, this peak extraction algorithm has a very important parameter, threshold, which should be set by the dynamic characteristics of the sound signal. The very peak will be selected which exceeds the threshold and is maximum in each windowed segment at the same time.

Since the peak selected by the peak extraction algorithm can reflect the intensity of the waveform change (slope and amplitude, called as the dynamical characteristics), we can get the RME temporal envelope by joining the adjacent peaks, only changing the peak value to a RME value. RME value is located somewhere between the peak value and the arithmetic mean, and shows a strong relative stability. The accuracy of extraction for the temporal envelope is depended on the width of the selected window. The smaller window width we use, the better transient response and more transient characteristics we will get. Peak extraction is a foundation of the temporal envelope, they are inseparably linked.

3.2 Dynamic analysis and its algorithm

In order to get the SPL dynamic of each sample, the original linear voltage value of the signals should be converted to the actual sound pressure level.

The sound pressure level L_p can be calculated as equation (1), with the sound pressure as p, and the reference sound pressure as p_0,

$$L_P = 10\lg\frac{p^2}{p_0^2} \tag{1}$$

Then, the SPL dynamic can be defined by the sound pressure level difference and converted to the sound pressure square ratio as equation (2).

$$L_p = L_{p_{\max}} - L_{p_{\min}} = 10\lg\left(\frac{p_{\max}^2}{p_0^2}\times\frac{p_0^2}{p_{\min}^2}\right) = 10\lg\frac{p_{\max}^2}{p_{\min}^2} \tag{2}$$

Set the binaural microphone sensitivity as S, in unit of V/Pa, the linear voltage signal is U, in unit of V, the actual sound pressure P corresponding to the microphone linear voltage can be calculated as equation (3).

$$p = \frac{U}{S} \tag{3}$$

At last, put equation (3) into equation (2), the signal dynamic can be calculated by the original linear voltage square ratio as equation (4):

$$L_p = 10\lg\frac{p_{\max}^2}{p_{\min}^2} = 10\lg\frac{(U_{\max}/S)^2}{(U_{\max}/S)^2} = 10\lg\frac{U_{\max}^2}{U_{\min}^2} \tag{4}$$

3.3 Extraction algorithms of IACF and IACC

Based on the characteristics of the human auditory system, Schroeder and others (1974) proposed a binaural evaluation parameter on concert hall sound quality evaluation (8). This parameter partial reflect the relative energy of early lateral reflections. The measurement use two small microphones which are located on the ear canal of human head or dummy head. The microphone’s electrical output (usually after recordings) is connected to the computer, then, the computer calculated the signal data with specific formula to get the value. Interaural cross-correlation coefficient (IACC, calculated from the interaural cross-correlation coefficient (IACF)) is such a parameter that measures the difference of the sounds between two ears at the same instant. Lateral sound into one ear earlier than into another ear, head will influence the sound propagation added, one sound will show difference features on two ears. If the sounds on two ears are totally different, IACC shows 0, which means the sounds on two ears are uncorrelated. On the contrary, the sound in front of you will bring identical sounds on two ears, IACC shows 1, which means no sense of space.

In this paper, IACC as equation (5) is a very important temporal feature, which can be used to analyze some soundscape characteristics related to the spatial awareness.

$$IACC = \max\left|IACF_{t_1 t_2}(\tau)\right| \tag{5}$$

For the binaural soundscape samples, IACF is defined as following:

$$IACF_{t_1 t_2}(\tau) = \frac{\int_{t_1}^{t_2} p_L(t)\, p_R(t+\tau)\,dt}{\sqrt{\int_{t_1}^{t_2} p_L^2(t)\,dt \cdot \int_{t_1}^{t_2} p_R^2(t)\,dt}} \tag{6}$$

The parameters, p_L and p_R denote the left ear and right ear signals. The maximum possible value of equation (5) is 1. The maximum time difference between two ears is around 1ms. Therefore, time delay τ is in the range of -1ms to 1ms.

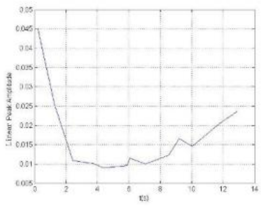
In this paper, all the binaural soundscape samples are digitized with sampling rate as 44100Hz. Then, the time delay τ is in the range of -44 samples to 44 samples. Since the aural sound image on left and right hemispheres are approximate symmetry, especially in the middle frequency and low frequency. Time delay τ is set in the range of 0 sample to 44 samples.

When analyzing the binaural soundscape samples in this paper, we do not only extract the IACC value, but also focus on IACF curve. At the same time, the mean and standard deviation values of IACF curve are also listed in table 2.

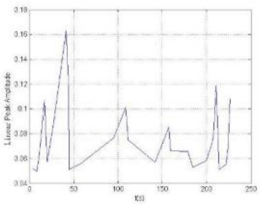
4. RESULTS AND ANALYSIS

4.1 RMS temporal envelope

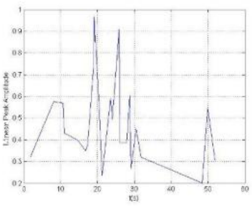
The temporal envelopes of the 22 binaural soundscape samples obtained by the algorithm of section 3.1 are shown in figure 3.



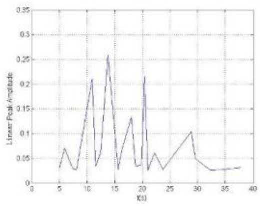
1



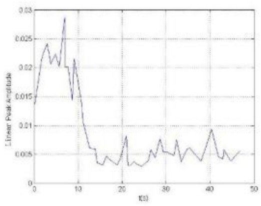
2



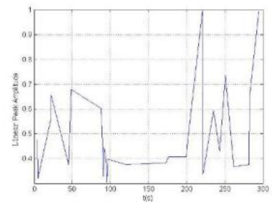
3



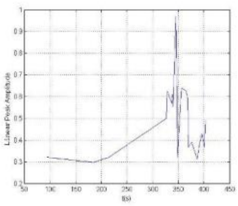
4



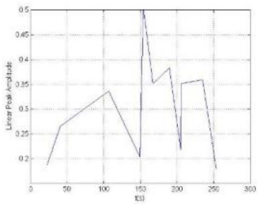
5



6



7



8

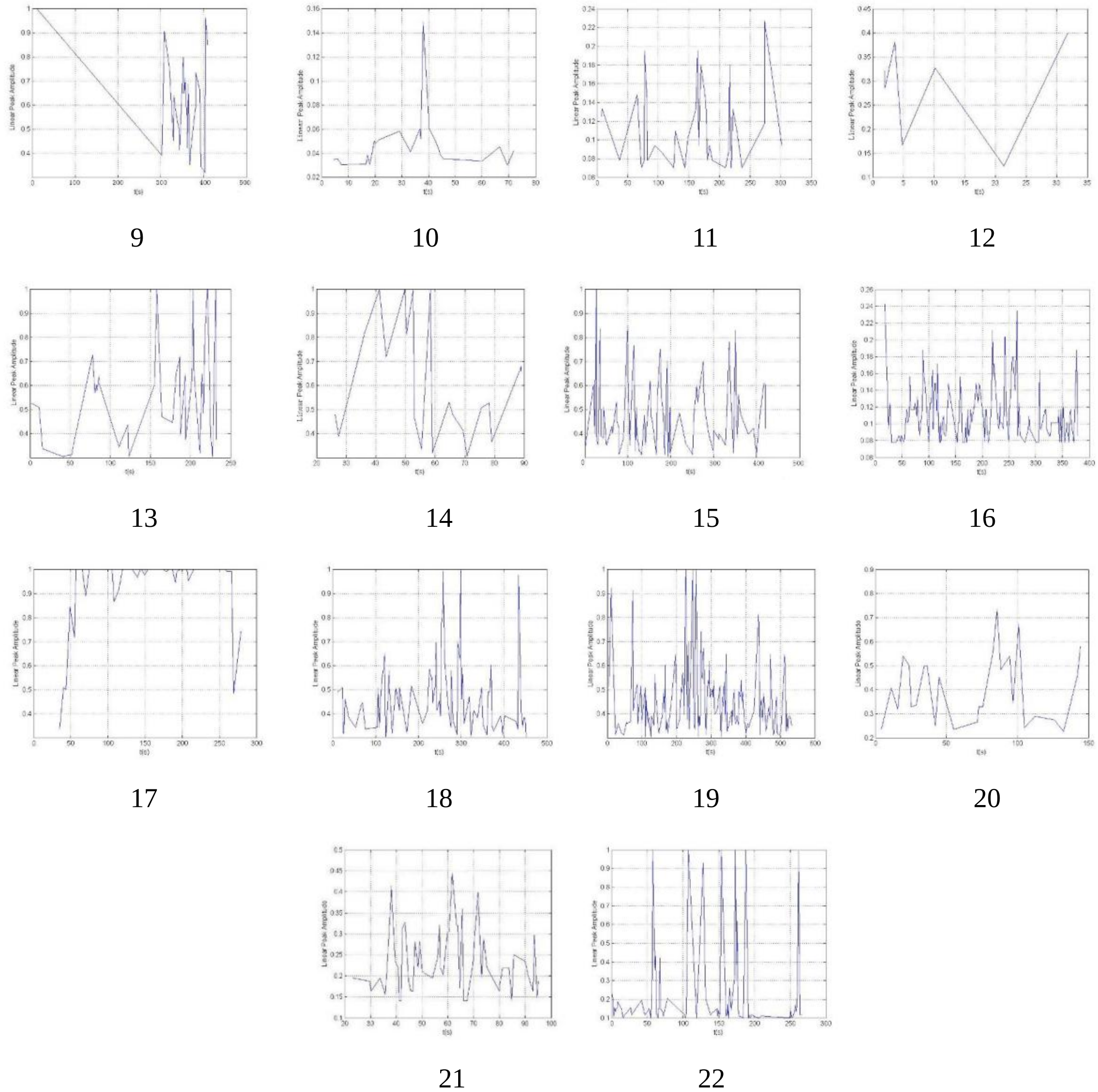


Figure 3 - The normalized RMS temporal envelope of the 22 binaural soundscape samples

The soundscape samples those collected at same location but different times may have some differences. Set sample 3 and 4 as example, they were collected at Madu village in the morning, but have subtle differences in the temporal envelope. The peaks of sample 3 are mainly concentrated in 5-12s and 16-31s, with sounds of birds and conversation, and the rest time is keynote with a lot of nature sounds. The peaks of sample 4 are mainly concentrated in 8-23s, with sounds of conversation from the locals. By comparing sample 3 and 4, we can find that samples collected from same location and nearly time still have subtle differences, especially in keynote, which can be intuitively manifested by temporal envelope. Temporal envelope not only able to showing the trends of the signal, but also able to express the content of the soundscape samples indirectly.

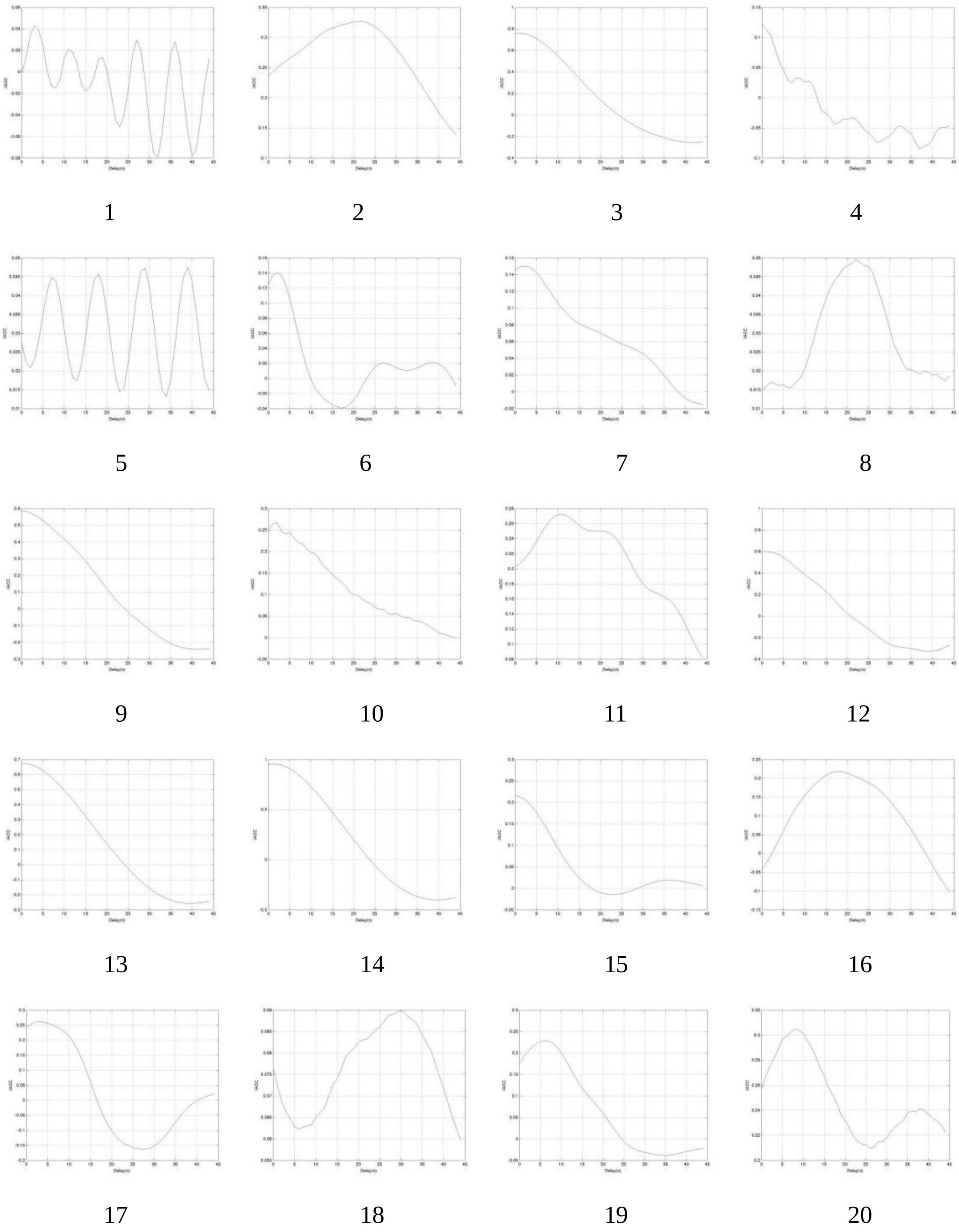
The soundscape samples from different locations have more differences. Set sample 14 and 20 as example, sample 14 was collected at Madu village, and sample 20 was collected at Yu-ma Park. The peaks of sample 14 are mainly concentrated in 35-58s, with soundmark of local pronunciation, and the rest time is keynote with a lot of natural sounds. The peaks of sample 20 are mainly concentrated in 15-47s and 80-105s, with soundmarks of peddler shouts and music, and he rest time is keynote with a lot of traffic noise. By comparing the sample 14 and 20, we can find that samples collected from different location have lots differences in keynote and soundmark, even in sense of hearing which can be reflected intuitively by temporal envelope. Soundscape samples of different place have it’s own characteristics. You can feel local culture and atmosphere through the soundscape, and that’s why soundscape is very charming.

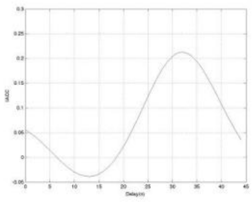
4.2 Soundscape dynamic analysis

The dynamics calculated by the method of section 3.2 are shown in table 2. We can get the conclusion that different soundscape samples may have different sound signal and keynote, which determine the dynamic of the sample. The dynamic of soundscape samples is determined by it's contents. By analyzing the relationship between the individual soundscape samples we can get that the soundscape samples which have different soundmarks may have large dynamic differences. Different collection locations and different sound contents may include different soundmarks, which determined the dynamic differences among soundscape samples.

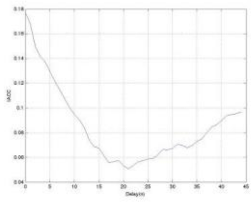
4.3 IACF and IACC analysis

IACF curves calculated by the method of section 3.3 are shown as figure 4, and their IACC values, IACF mean values and the standard deviations are shown in table 2.





21



22

Figure 4 - IACF curves of the 22 binaural soundscape samples

Table 2 – IACC and dynamics of the 22 binaural soundscape samples

No.	IACC	Mean of IACF	Standard Deviation	L_{eq} (dBA)	Dynamic (dB)
1	0.04	-0.010	0.033	48	14
2	0.33	0.261	0.054	60	10
3	0.76	0.160	0.370	76	13
4	0.12	-0.020	0.055	55	19
5	0.05	0.030	0.011	40	21
6	0.14	0.020	0.050	72	26
7	0.15	0.065	0.051	70	10
8	0.05	0.028	0.012	57	12
9	0.59	0.111	0.297	70	20
10	0.27	0.111	0.085	50	14
11	0.27	0.203	0.054	55	8
12	0.60	0.047	0.338	69	19
13	0.67	0.132	0.344	75	26
14	0.96	0.180	0.510	75	22
15	0.22	0.048	0.073	79	30
16	0.22	0.102	0.098	65	10
17	0.26	0.022	0.155	85	15
18	0.09	0.074	0.010	73	26
19	0.23	0.066	0.103	75	34
20	0.31	0.244	0.031	74	9
21	0.21	0.070	0.085	69	10
22	0.18	0.084	0.031	65	20

Then, clustering the 22 samples according to the IACC values, stepping as 0.1, the rough relationship between the average L_{eq} of each cluster range and corresponding IACC values is shown in table 3. There are positive correlation between IACC and L_{eq}, except the range of 0.5-0.6 and 0.9-1.0. But there are only one sample in the range of 0.5-0.6 and 0.9-1.0, and one sample can not represent the average L_{eq} at all.

Table 3 - Clustering of soundscape samples

IACC	Quantity	Number of the samples	Average L_{eq} (dBA)
0-0.1	4	1、5、8、18	54.57
0.1-0.2	4	4、6、7、22	65.66
0.2-0.3	7	10、11、15、16、17、19、21	68.21
0.3-0.4	1	2、20	66.87
0.4-0.5	0	--	--
0.5-0.6	1	9	69.94
0.6-0.7	2	12、13	72.1
0.7-0.8	1	3	75.57
0.8-0.9	0	--	--
0.9-1.0	1	14	74.92

5. CONCLUSION

Soundscape contains pluralistic ideology and is involved in many fields, such as noise control, subjective evaluation of soundscape, soundscape design and soundscape composition. The diversified and variability content of soundscape causes the diversity of the soundscape. Geographical and cultural characteristics make the soundscape totally different from other types of music.

Soundscape waveform can be represented by RMS temporal envelope, and the structure of soundscape temporal envelope indirectly reflects the soundscape contents. There are two main aspects differences among soundscape samples: time differences and geographical differences. Time difference performed as samples collected at same place, and have same soundmarks, but have different sound signals and keynotes. Geographical differences performed as samples with a lot of differences in soundmarks, sound signals and keynotes. As precious binaural information, soundscape contents can reflect regional culture and customs. Then, dynamics of soundscape samples can show dynamic differences of each sample itself and differences among samples. The interaction among soundmarks, sound signals and keynotes determines the dynamic of soundscape samples. Finally, there are positive correlation between IACC and equivalent sound pressure level of soundscape samples. IACC and IACF maybe considered as some of objective parameters to describe subjective special awareness, but still need a further study.

REFERENCES

1. R. M. Schafer. The Soundscape:Our Sonic Environment and the Tuning of the World. New York, USA: Alfred Knopf; 1997.
2. Jian KANG. Urban Sound Environment [Chinese Edition]. Beijing, China: Science Press; 2011.
3. Jun YAN, Gary. Sound and the City [Chinese Edition]. Shanghai, China: Shanghai Renmin Press, 2007.
4. Deng Zhiyong, Wang Daiwei, Liu Aili, Chen Haokui, Zhang Jian. Semantic Assessment for the Soundscape of Chinese Ethnic Historical Areas. Proc INTER-NOISE 2013; 15-18 September 2013; Innsbruck, Austria 2013. No.0123, p.268.
5. Zhiyong Deng, Guowen Zhou, Hua Wei, Yun Wang, Jian Zhang. Soundscape Case Study: Acoustics, Ecology and its Anthropological Sense. Proc ACOUSTICS 2012; 13-18 May 2012; HK, China 2012. p. 290.
6. Yun Wang, Zhiyong Deng. Soundscape: In the View of Music. Proc INTERNOISE 2011; 4-7 Sep. 2011, Osaka, Japan 2011. No. in11_427903, p.3934-3939.
7. YAO Shun, LIU Hai-long, CHEN Chuan-ping, LI Xiang-ning. An Improved Approach of Peak – detection Algorithm for Spike Acquiring. Journal of Biomedical Engineering Research [Chinese Edition]. 2005; Vol.1:14-17.
8. Jens Blauert. Communication Acoustics. Berlin Heidelberg, Germany: Springer; 2005.