CLASSIFICATION OF SOUNDSCAPES BASED ON THEIR DYNAMICS

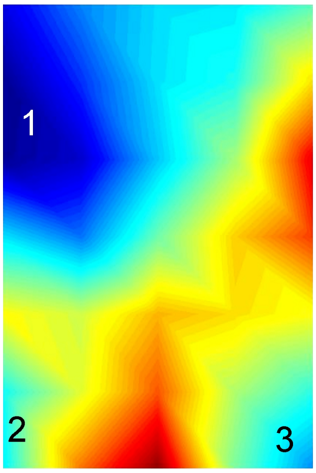
B. De Coensel, D. Botteldooren, T. De Muer
Acoustics Group, Department of Information Technology, Ghent University, Belgium

Introduction Soundscape research extends the range of effects of noise on man that are considered as well as the number of non-acoustic factors that are included in impact research. Classification and characterisation are often a first step, so several methodologies were proposed for this purpose [1]. Perception and quality studies [2, 3, 4] have identified a factor that relates to dynamics of the sound, hence the aim of this paper to use it for classification.
The long-term variations in level and pitch of different kinds of music mankind has produced have already been studied in the late 1970s [5]. All of the musical genres that were studied showed the same *1/f* spectral behaviour both for level and pitch variations. Artificial sound with these characteristics was also recognized as "music" by a listener. Sound with a flatter spectrum sounded too chaotic, too unpredictable. A steeper slope resulted in too much predictability and hence boring and dull sound. These findings were so surprising that they have puzzled many. In 1987 however, Bak et al. [6] proposed that many complex systems naturally self-organize to a critical state, with the consequent scale-free fluctuations giving rise to spectral power laws. In creating music, man thus seems to imitate the temporal fluctuation of self-organized critical systems, which are quite common in (natural) living environments.
Given the above observations it seemed obvious to analyse the spectrum of amplitude and pitch fluctuations of urban and rural soundscapes and to look for *1/f*-like features. At one hand, the urban and rural soundscape can be assumed to be the voice of a complex system. At the other hand similarity to music can be an indication of quality.

Methods Sound fragments of 15 minutes were used for this study. These were recorded monaurally using an omni-directional microphone. The signal processing presented in [5] was used, although A-weighting replaced audio range bandpass filtering. After applying a digital A-weighting filter, the signals were squared. Hereupon a 1st order digital Butterworth lowpass filter with cut-off frequency 20 Hz (roughly corresponding to a 50ms averaging time) was applied. The result is the instantaneous A-weighted signal energy. The spectral density of A-weighted level variations was calculated using the standard FFT scheme after down sampling, using a rectangular time window of length 15 minutes. Finally, the curves obtained in the log(amplitude) versus log(frequency) domain were smoothed by 1/24-octave band averaging.
For urban and rural soundscapes, often a breakpoint seems to occur in the spectra between 0.1 Hz and 1 Hz. In an attempt to reduce data for further analyses, the frequency interval is therefore split into 2 parts: I_1 = [0.002Hz, 0.2Hz] and I_2 = [0.2Hz, 5Hz]; $I_3 = I_1 + I_2$ is the full frequency range. I_2 corresponds to a time interval between 200ms and 5s and can be seen as characteristic for the sound fluctuations emerging from the source itself. I_1 corresponds to the interval between 5s and about 10min. It is influenced mainly by series of events such as cars, trains, planes, or talking people passing the observer. In each interval, the slope α of a linear fit on the log-log chart is calculated together with the squared fitting error ε. This results in 6 indicators for the soundscape dynamics. A classification method is proposed based on self-organizing maps (SOM's), a type of unsupervised learning neural networks introduced by T. Kohonen [7]. In particular, the SOM Toolbox [8] for Matlab was used.

Results and discussion A dozen of recordings were made in rural (R) area in Flanders; additionally 15 urban (U) soundscapes were recorded in the city of Ghent, Belgium [9]. For comparison, also music (M), speech (S), some typical "natural" sounds (N, fountain, dawn

N4	M11	S U9	N1 N2	W
M1 M3	M10		M12 M15	U5 M5
U11 M14	M2 M13	U4 M4	M6	
R6	U6 R1		R4	U13 M8
R8		N3 R5	U15	U2 U10
M9 R3	U14			U3 U7 U12
U1 M7 R7		U8 R11	R12	R2 R9 R10



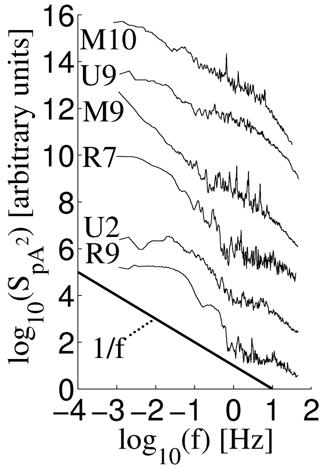


Figure 1, 2: Self-organizing map & accompanying unified distance matrix. *Figure 3:* Power spectrum of A-weighted level variations of (M9,10) music of M.Ravel & J.S.Bach, (U9) traffic free shopping street, (R7,9) typical rural soundscapes in Flanders, (U2) narrow street canyon.

chorus of birds, waves at rocky coast) and white noise (W) were included in the classification. Figure 1 shows a SOM, where the labels of the input data are placed at the best matching neuron. Fragments at the same neuron thereby roughly have the same characteristics (e.g. the speech fragment S and the car free shopping street U9). The difference in characteristics between neighbouring neurons can be visualized in a unified distance matrix as seen in figure 2, where red means large difference. Three clusters located at the blue valleys can be identified; two spectra from each cluster are drawn in figure 3.

Cluster 1 corresponds with approximate *1/f* dynamics ($\alpha \approx -1$) and relative small deviations from the power law (ε small) for the whole frequency range, typical for a complex system. Most of the music and natural sounds are located in this cluster, together with most of the urban fragments wherein the murmur of passing pedestrians dominates (U4,6,9,11). Fragments in cluster 2 show a steeper slope with very small ε in I_1, and *1/f* behaviour in I_2, but with large deviation ε. This implies some predictability on longer timescales. The musical fragments located in this cluster illustrate this: Barber's *Adagio* (M7) and Ravel's *Bolero* (M9), both pieces consisting of one long musical sentence. Cluster 3 combines fragments with spectrum slopes steeper than *1/f* and with large ε. This implies more slow variations than expected in the case of SOC. Single cars and planes passing by and heard from far away may cause this.

Conclusion The spectrum of noise level fluctuations and in particular its linearity on a log-log scale can be used to classify soundscapes on the bases of the long-term dynamics. This classification shows surprising similarities between natural, rural, and urban soundscapes and music that could become the object of future soundscape research.

Keywords: soundscape, *1/f* noise, self-organized criticality, self-organizing map.

References

1. B. Schulte-Fortkamp, proceedings of Forum Acusticum, Sevilla, Spain, 2003.
2. S. Viollon and C. Lavandier, proceedings Internoise 2000 (CDROM), Nice France, 2000.
3. V. Maffiolo et al., proceedings Internoise 1999, Fort Lauderdale, Florida, USA, 1999.
4. B. Berglund and M. Nilsson, NOPHER 2001, Cambridge, UK, 2001.
5. R. Voss and J. Clarke, J. Acoust. Soc. Am., 1978.
6. P. Bak et al., Phys. Rev. Lett., 1987.
7. T. Kohonen. *Self-organizing Maps*. Springer Verlag, Berlin, 1995.
8. http://www.cis.hut.fi/projects/somtoolbox/
9. Locations of the recordings, together with some sample fragments, can be found at http://www.intec.rug.ac.be/groupsites/acoustics/GentSoundscapes.html