ACOUSTICS 2012
HONG KONG

The role of paying attention to sounds in soundscape perception

Dick Botteldooren[1], Michiel Boes[1], Damiano Oldoni[1], Bert De Coensel[1]

[1] Acoustics Group of Department of Information Technology,

Ghent University, Gent, Belgium

Abstract: It has been stated frequently that the soundscape as perceived and appraised by the user of a space, extends beyond the physical stimulus. We argue that, when introducing to human-factor in analyzing a sonic environment, the sounds that people hear play an important role. This holds in particular for rather quiet and infrequent disturbance of park soundscapes. Auditory attention mechanisms are essential in the process. Attention can be drawn by saliency elements such as changes in time and frequency, but it can also be outward oriented and voluntary. These mechanisms could explain the special role of natural sounds in distracting attention from mechanical background hum in a park environment. These theoretical concepts have now been implemented in measuring equipment that allows estimating how often particular sounds will be heard by a human listener. The methodology includes biologically inspired feature extraction, learning based on co-occurrence of features and saliency, attention focusing, and inhibition of return. Extension to binaural measurements increasing the unmasking effect is also discussed.
Keywords: Soundscape, Auditory attention,

1. Introduction

The study and design of urban sonic environments has since a few decades embraced the soundscape concept. Nevertheless very little has been achieved regarding detailed theoretical understanding, modeling and measuring. Quite often, researchers and designers turn to classical L_{eq} based noise maps and level measurements when bringing their ideas to practice. New theories, models and measurement techniques are needed. These should relate the physical sounds that are present in the sonic environment to the sound objects that are the more important building blocks of the soundscape object in the human mind. Attention to those physical sounds plays an important role in determining the importance of them in the soundscape. In this paper we focus on this role of attention.

In Section 2 a theory for soundscape is further developed to include knowledge on attention mechanisms gathered from psychology and physiology, in particular neurology. This theory is used in Section 3 to explain known dependencies in soundscape research. Section 4 explains how attention mechanisms could be included in measurement instrumentation.

2. A soundscape theory

In a strict sense soundscape could be regarded as the object existing in the observers mind to stress the importance of the listener. However, by extension authors often refer to the physical environment as the soundscape. In this section we adhere to the first approach. The purpose of the proposed soundscape theory is to gather information from different disciplines regarding the perception of sound and environmental sound in particular and present it in a concise and self consistent way. This allows combining the growing insight provided by neuroscience and in particular brain imaging with psycho-physics knowledge.

In a previous publication [Botteldooren et al., 2011] a theory on soundscape obtained from the work of COST TD0804 was presented. Different styles of listening were identified as the first step in constructing the soundscape in the listeners mind. As the listening experience in a sonic environment evolves, the listener switches between different listening styles: from the more holistic *listening in readiness* waiting for familiar or important sounds to emerge (expected or not), to *listening in search* expecting particular sounds in a context, or even to *story listening* focusing attention on one particular sonic story within the multitude of sounds. Listening as a whole can be using more or less cognitive processing resources leading to foregrounding (attentive listening) and backgrounding (holistic listening) [Truax, 2001]. Switching between these types of listening depends on multi-leveled attention processes and higher cognitive functions.

Let us now focus on the attention mechanism in more detail. One could say that main role of selective attention is to prevent sensory signals from overloading the higher level cognitive system. For experiencing the sonic environment, not attentively listening to the sounds is the default state. However, as the auditory system always stays alert, sounds within the sonic environment could draw attention. The proposed theoretical model foresees a two stage mechanism to account for this: auditory stimuli draw attention because of specific features they possess but they don't necessarily get attended to. This two stage mechanism is supported by neuroscience: sounds with high saliency trigger early brain response [Escera et al., 1998] while inhibition of return [Prime et al., 2003] and general attentiveness to sound determine whether a late response corresponding to actual attending is observed. Recognizing sound features that increase saliency [Kayser et al 2005] and attract attention is an important aspect of the proposed soundscape theory. It is well known [Kayser et al 2005] that spectral and temporal variations and modulations – sometimes referred to as ripple – increase saliency for human observers. However, the auditory brainstem, which is responsible for these specific sensitivities, has a much higher plasticity than one might expect. A well known example of this can be found in the experiments with rats by Chang and Merzenich (2003). The importance of speech (and music) for humans is expected to shape the sensitivity for spectro-temporal changes in early childhood. On the basis of this, one could expect a common basis for auditory saliency, but in addition some specificity for different (groups of) people.

Auditory streams are classically regarded as existing in a pre-attentive phase. Although this view is appealing because of its conceptual simplicity, recent findings suggest that attention also plays a role in the formation of auditory streams [Cusack et al. 2004][Shamma et al. 2011].Overall, it can be stated that the process of auditory scene analysis draws on low-level principles for segmentation and grouping, but is fine-tuned by selective attention [Fritz et al., 2007]. Sound objects within the sonic environment are thus formed with the help of selective attention.

The listener embedded in a real environment – in contrast to experimental conditions – relies on all senses to structure the image of the environment in its mind [Driver and Spence, 1998]. One sensory modality could draw spatial attention also to a different modality and even influence the perception itself strongly. This raises the question whether attention resources are controlled by a supramodal system or by many modality specific attention systems. In focused attention conditions, judging each signal (sound and vision) separately when incongruent sensory signals occur at the same location is difficult, at least much more difficult than when the incongruent signals come from different spatial location and attention is divided [Santangelo et al., 2010]. A multilevel mechanism of attention with a multimodal component overarching the single sensory component seems the most plausible model given today's knowledge. In the context of assessing the sonic environment, this could be interpreted as a stronger emphasis on visual sources but at the same time a lower identification of deviant sound experience if it comes from the same location.

Listening in search or story listening involves voluntary (endogeneous) attention focusing grounded in higher level cognition. It can be shaped by expectations about the place based on prior experience or knowledge or it can be initially triggered by involuntary attention focusing. In the latter case, incongruence of the sound in the scene can enhance detectability [Gygi and Saphiro, 2011]. Event Related Potential measurements confirm the deviant processing also with complex sounds but also show that familiarity with the sound has an effect [Kirmse et al. 2009]. A foundation for rapid extraction of meaning from a familiar environmental sound was observed even when sounds were not consciously attended. Thus the soundscape theory has to account for this dual effect: congruent and familiar sounds are less likely to trigger attention but they are also to most probably object of voluntary attention focusing during listening is search or story listening.

Until now we did not consider the relationship between attention and binaural hearing. Inhibition of return on location [Mondor et al., 1998] could explain why moving sources or groups of sources of the same kind popping up at different locations might be less easily inhibited by the auditory system and thus continue to attract attention longer than a stationary source. It is known that identity information predominates over location information in auditory memory [Mayr et al., 2011] thus soundscape appraisal in itself – in contrast to unmasking – may be less sensitive to aspects of binaural hearing.

The reaction of the brain to sensory stimuli depends on its current state. According to the attention to memory model hypothesis that very similar attention mechanism are involved in memory tasks at the one hand and sensory processing tasks at the other [Cabeza et al., 2011]. Part of the neural circuitry even seems to overlap. This implies additional modulation of overall attention devoted to the sonic environment. Conversely, it also implies that sensory input in general and sound in particular can distract from memory (and cognitive) tasks. Soundscape perception can therefore be different for the same person at different instances. Internal variables in attention models need to account for this.

3. Explaining known dependencies based on attention mechanisms in the soundscape theory

Attention mechanisms can – at least partly – explain several of the observations made with regard to soundscapes and their effects on people by several authors. As part of a soundscape design process, several authors have pointed out that adding sounds that are appraised by a large part of the population as a pleasant and stimulating component of the sonic environment can improve the overall soundscape without lowering the level of unwanted sound. Water sounds have been studied extensively e.g. [Jeon et al., 2010] for this purpose. Similarly, bird song [De Coensel et al., 2011], gentle voices, wind, etc. have been shown to improve the perceived quality of a sonic environment. Comparing the level and spectrum of the wanted sound with the unwanted sound, it is clear that inner ear masking effects cannot account for these observations, in particular when the wanted sound has strongly fluctuating amplitude. Attention mechanisms explained above can easily explain these observations: vocalizations attract attention because of their saliency and thus distract attention from unwanted sounds. It is perfectly clear from the discussion in the previous section that vocalizations should indeed be observed more easily because of their relevance for human communication.

The importance of the interaction between aural and visual stimuli for soundscapes has been investigated and stressed by several authors, e.g. [Yang and Kang, 2005][Raimbault and Dubois, 2005][Viollon et al., 2002]. The multimodal component in the attention mechanism can explain many of these observations. Visual characteristics of an object, e.g. its incongruence in the visual scene or simply the fact that the object moves can attract visual attention and thus also focus auditory attention on this object. One could be tempted to deduct that invisible sound sources are less disturbing, but this is not necessarily through. If the sound still captures attention although it is invisible, its unexpected nature or incongruence with the visual scene may trigger higher cognitive mechanisms leading to negative appraisal.

Cultural and social differences have been observed in the way the sonic environment is perceived and appreciated by persons using this environment. The different meaning given to the sounds by people with different socio-cultural background is an obvious candidate for explaining these differences. However, attention could also play a (small) role here. The plasticity of the auditory brainstem [Bidelman et al., 2011] could make particular groups of people more susceptible and sensitive to particular sounds thus leading to stronger sensory driven attention for particular types of sounds. Once a sound is noticed, meaning could help to trigger more outward oriented attention, leading to strong focusing on for example unwanted components of the sonic environment and subsequent negative appraisal.

The potential for mental restoration of the public open space in urban environments is one of the main reasons for promoting conservation of than public space. It has been suggested frequently that a matching soundscape could enhance the potential for mental restoration of such a space, but the experimental observation of this effect is rare. In lab environment, researchers at Stockholm University clearly showed the advantage of natural sound mixes over noise at different levels for recovery from psychological stress [Alvarsson et al., 2010]. The authors argue that "the mechanisms behind the faster recovery could be related to positive emotions (pleasantness), evoked by the nature sound …". The experiment also included quiet ambient sound which turned out to have less restorative power. The authors offer the lack of information in the sound as a reason for this. The latter explanation could be refined by the attention mechanisms discussed above. In line with [Kaplan, 1995], psychological restoration requires that internally directed attention is disrupted by external stimuli. A soundscape rich in variety of sounds that easily attract human attention would therefore be beneficial for psychological restoration. The role of pleasantness may to some extent be explained through enhanced cognitive attention for the sound.

4. Bringing theory to practice: measuring

Bringing the theory for soundscape to measurement equipment requires several simplifications. Features extracted from the incoming sound stream are to be calculated on equipment with limited computational capabilities or should be based on a limited number of quantities to be transmitted from the sensor node to a computational node located elsewhere on the internet. Spectral resolution was therefore limited to one-third-octave bands. Loudness is calculated to account for cochlear mechanisms including energetic masking. Features reflecting temporal and spectral contrast are extracted using double Gaussian filters. Spectral harmonics and rhythm are not explicitly taken into account. Details can be found in [Oldoni et al., 2010].

The computational model has to tackle the problem of sound identification and stream segregation. Inspired by human auditory processing, the auditory feature space is mapped to a two dimensional space by a self organized mapping (SOM) that relies on co-occurrence (temporal coherence) of features to introduce structure. As the mapping needs more specificity where the human listener would be able to discriminate between sounds, training should be enhanced for sounds that attract attention. A short-cut consists in using saliency as a proxy for bottom-up attention mechanisms. The training of the map is continued as long as the measurement system is placed in a sonic environment. This reflects the plasticity of the human auditory processing. New sounds occurring in a place will initially be treated as novel and unknown but will gradually become more familiar.

All sound in the sonic environment being represented as efficiently as possible – with respect to human-like capabilities – in a two-dimensional map, attention mechanisms are implemented as a spotlight on this map. Attention is modeled as a result of the interplay between activation and inhibition of all the nodes in the two-dimensional map. Activation is caused by the sensory stimuli and their saliency at the one hand and by lateral activation at the other. The latter is needed to create congruent areas in the map to receive the spotlight jointly. Two forms of inhibition are implemented in the model. Inhibition of return – preventing focusing attention for too long on a sound – grows steadily as a node in the map receives attention and then very slowly fades away after the node was no longer attended to. The second form of inhibition prevents all the nodes in the map but a few to be selected by the attention mechanism. It enhances contrast. Within the context of the model it can be called global inhibition since it affects the whole map. Voluntary attention leading to listening in search or story listening cannot be predicted as such, but its effect on gating stimuli to working memory can. The effect of this form of attention could be implemented as an additional excitation of the parts of the map corresponding to the target sound combined with a reduction of inhibition of return mechanisms. The former could be seen as mimicking the formation of a mental image of a sound that the human listener is searching for, the latter mimics the continuing interest in the sound [Boes et al. 2012].

Models for auditory attention embedded in soundscape measuring equipment cannot account for attention focusing on other sensory input nor on internally oriented attention. Yet, as was discussed above, biological attention mechanisms are at least partly shared by different senses. If the computational model neglects to account for this, unexpected emerging behavior is observed: even during quiet periods, the system tries to focus on a sound. Thus a simple statistical approach to other focus of attention has to be added.

As an example the result of running this computational model for soundscape perception on the signal from one of the urban microphones placed in the city of Gent. Figure 1 shows the frequency with which a variety of sounds – mapped in a two dimensional map – is expected to be noticed or paid attention to by the average listener.

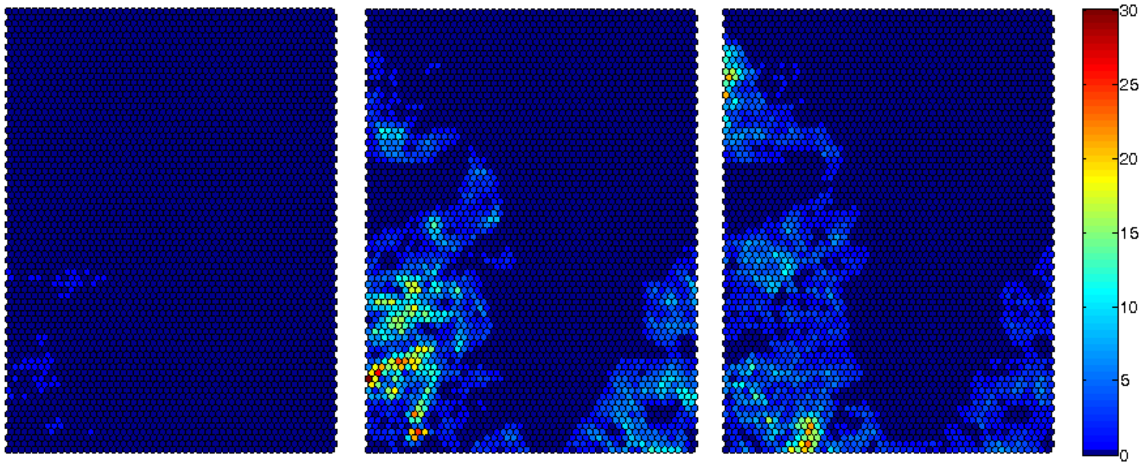

Figure 1 – Frequency of noticing or attending to different sounds: left 1am-2am, middle 7am-8am, right 1pm-2pm. At lower left in the map individual car passages are found, at upper left construction noises, upper right represents city background sound, etc. These sounds can be discovered by clicking on the map.

5. Conclusions

A theory explaining the important role of attention in soundscape research and design is proposed in this paper. It is shown how this theory can explain several observations made by soundscape researchers. A practical application of the insights gained from formally describing how attention might influence the formation of a soundscape object is illustrated by presenting an implementation in measurement equipment.

Acknowledgement

Michiel Boes is a doctoral fellow, and Bert De Coensel is a postdoctoral fellow of the Research Foundation–Flanders (FWO–Vlaanderen); the support of this organisation is gratefully acknowledged.

References

Alvarsson, J. J., Wiens, S., and Nilsson, M. E. (2010). "Stress Recovery during Exposure to Nature Sound and Environmental Noise," International Journal of Environmental Research and Public Health 7, 1036-1046.

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). “Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem.” Journal of Cognitive Neuroscience, 23, 425–434.

Boes M., Oldoni D., De Coensel B., and Botteldooren D., “Attention-Driven Auditory Stream Segregation using a SOM coupled with an Excitatory-Inhibitory ANN,” accepted for Proceedings of the 2012 IEEE World Congress of Computational Intelligence (WCCI), Brisbane, Australia.

Botteldooren D., Lavandier C., Preis A., Dubois D., Aspuru I., Guastavino C., Brown L., Nilsson M. & Andringa T.C. (2011) Understanding urban and natural soundscapes. Proceedings of Forum Acusticum, Aalborg, Denmark.

Cabeza R., Mazuz Y.S., Stokes J., Kragel J.E., Woldorff M.G., Ciaramelli E., Olson I.R., and Moscovitch M., (2011) “Overlapping Parietal Activity in Memory and Perception: Evidence for the Attention to Memory Model,” Journal of Cognitive Neuroscience 23:11, 3209-3217

Chang EF and Merzenich MM. (2003) “Environmental noise retards auditory cortical development.” Science 300: 498–502.

Cusack R., Decks J., Aikman G., and Carlyon R.P. (2004), “Effects of location, frequency region, and time course of selective attention on auditory scene analysis”, J. Exp. Psychol.–Hum. Percept. Perform. 30, 643–656.
De Coensel B., Vanwetswinkel S., and Botteldooren D. (2011), “Effects of natural sounds on the perception of road traffic noisen“ J. Acoust. Soc. Am. 129, EL148, DOI:10.1121/1.3567073
Driver J., Spence C. (1998), “Attention and the crossmodal construction of space”, Trends in Cognitive Sciences, Volume 2, Issue 7, Pages 254-262, 10.1016/S1364-6613(98)01188-7.
Escera C., Alho K., Winkler I., and Nätänen R., (1998). “Neural mechanisms of involuntary attention to acoustic novelty and change”, J. Cogn. Neurosci. 10, 590–604
Fritz J.B., Elhilali M., David S. V., and Shamma S. A. (2007), “Auditory attention — focusing the searchlight on sound”, Curr. Opin. Neurobiol. 17, 437–455.
Gygi, B., and Shafiro, V. (2011). "The Incongruency Advantage for Environmental Sounds Presented in Natural Auditory Scenes," Journal of Experimental Psychology-Human Perception and Performance 37, 551-565.
Jeon J. Y., P. J. Lee, J. You, and J. Kang, (2010) “Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds,” J. Acoust. Soc. Am., vol. 127, no. 3, pp. 1357–1366
Kaplan, S (1995). The restorative benefits of nature—toward an integrative framework. J. Environ. Psychol. 1995, 15, 169-182.
Kayser C., Petkov C., Lippert M., and Logothetis N. K., (2005) “Mechanisms for allocating auditory attention: An auditory saliency map”, Curr. Biol. 15, 1943–1947
Kirmse U., Jacobsen T., Schröger E. (2009), “Familiarity affects environmental sound processing outside the focus of attention: An event-related potential study,” Clinical Neurophysiology, 120, Issue 5, 887-896, ISSN 1388-2457, 10.1016/j.clinph.2009.02.159.
M. Elhilali, J. Xiang, S. A. Shamma, and J. Z. Simon, “Interaction between attention and bottom-up saliency mediates the representation of foreground and background in an auditory scene”, PLoS Biol. 7, e1000129 (2009).
Mayr, S., Buchner, A., Moller, M., and Hauke, R. (2011). "Spatial and identity negative priming in audition: Evidence of feature binding in auditory spatial memory," Attention Perception & Psychophysics 73, 1710-1732.
Mondor, T. A., Breau, L. M., and Milliken, B. (1998). "Inhibitory processes in auditory selective attention: Evidence of location-based and frequency-based inhibition of return," Perception & Psychophysics 60, 296-302.
Oldoni, D., De Coensel, B., Rademaker, M., De Baets, B., and Botteldooren, D. (2010). "Context-dependent environmental sound monitoring using SOM coupled with LEGION," Proceedings of the 2010 IEEE World Congress of Computational Intelligence (WCCI), Barcelona, Spain.
Pheasant R. J., K. V. Horoshenkov, G. R. Watts, B. T. Barrett, (2008) “The acoustic and visual factors influencing the construction of tranquil space in urban and rural environments: – Quiet Places?” Journal of the Acoustical Society of America, 123, 1446 – 1457
Prime D. J., Tata M. S., and Ward L. M. (2003), “Event-related potential evidence for attentional inhibition of return in audition”, NeuroReport 14, 393–397
Raimbault M., Dubois D., (2005) “Urban soundscapes: Experiences and knowledge”, Cities, 22, Issue 5, 339-350
Santangelo V., Fagioli S., Macaluso E. (2010), “The costs of monitoring simultaneously two sensory modalities decrease when dividing attention in space”, NeuroImage, Volume 49, Issue 3, Pages 2717-2727, 10.1016/j.neuroimage.2009.10.061.
Shamma S. A., Elhilali M., and Micheyl C., (2011) “Temporal coherence and attention in auditory scene analysis”, Trends Neurosci. 34, 114–123
Truax, B. (2001), Acoustic communication (2nd ed.). Westport, CT, Ablex pub.
Viollon S., Lavandier C., Drake C. (2002), “Influence of visual setting on sound ratings in an urban environment”, Applied Acoustics 63, Issue 5, 493-511, 10.1016/S0003-682X(01)00053-6.
Yang W. and Kang J., “Acoustic comfort evaluation in urban open public spaces”, Appl. Acoust., 66, 211–229, (2005).