Sound Ambiences Consistent with Visualizations of Landscape Types: An Evaluation

Ulrike Wissen Hayek[1], Laura Endres[2], Reto Spielhofer[2], Adrienne Grêt-Regamey[2]

[1]Planning of Landscape and Urban Systems, ETH Zürich, Zürich/Switzerland· wissen@nsl.ethz.ch
[2]Planning of Landscape and Urban Systems, ETH Zürich, Zürich/Switzerland

Abstract: Using digital landscape representations for assessing people's perceptions of the visual landscape and their preferences for alternative options of landscape change, a high simulation quality is required. Thereby, sounds can augment people's experiencing of the demonstrated landscape. We present an approach to establish sound ambiences and to investigate their level of perceived consistency with visual representations of different landscape types in a laboratory experiment. The results show that sounds expected through the visual contents and sounds anticipated by participants remembering similar situations need to be reproduced. For presenting a realistic sound ambience a mix of soundmarks and of more general ambient sounds is important. Further, the sounds' volume and dominance are decisive for the overall consistency. It is notable that people in laboratory environments seem to not always accept the volume measured in the real landscape as appropriate. Our conclusions give advice for designing audio-visual simulations. Furthermore, we provide guidance for evaluating the consistency of audio-visual stimuli for implementation in landscape preference studies.

Keywords: Audio-visual simulation, environmental perception, landscape types, perception study

1 Introduction

Landscape typologies, i. e., systematic classifications of generic landscape types by defined attributes, are a valuable basis for landscape assessment on a (supra-)national scale and for assigning policy response (VAN EETVELDE & ANTROP 2009 VAN DER ZANDEN et al. 2016, ARE, BAFU, BFS 2011). In Switzerland, for example, there is a need for a spatial coordination of tasks to plan a mix of renewable energy infrastructures and their prioritization in the various landscape types. In this context, public judgments of the landscape impacts stimulated by a mix of renewable energy systems in Swiss landscapes will be assessed to recommend a prioritization of such systems.

To improve landscape development concepts, people's perceptions of the visual landscape and their preferences for alternative options of landscape change are assessed. Thereby, visual representations of the scenarios proved to be effective media to illustrate possible landscape changes. They trigger affective responses regarding visual landscape preferences, but a high simulation quality is required (VAN BERKEL & VERBURG 2014, CELIO et al. 2015, RIBE et al. 2017). Not only is the visual realism of 3D landscape simulations important. Studies have shown that environmental sounds can increase the vividness and experimental authenticity of simulated landscapes and intensify the study participants' immersion into the virtual landscape scenes (LINQUIST et al. 2016, WISSEN HAYEK et al. 2016). The question remains, which environmental sounds are required and how they should be coupled with the representations of generic landscape types.

Landscape typologies are predominantly based on natural features whereas perceptual properties play a minor role in characterizing the landscape types (VAN EETVELDE & ANTROP 2009). However, a multi-sensory environmental characterization is necessary to fully assess

Journal of Digital Landscape Architecture, 3-2018, pp. 172-180. © Wichmann Verlag, VDE VERLAG GMBH · Berlin · Offenbach. ISBN 978-3-87907-642-0, ISSN 2367-4253, e-ISSN 2511-624X, doi:10.14627/537642018.

the character of landscape types (SWANWICK 2002). In particular, including environmental sound into the assessment is increasingly gaining attention. Efforts are made, e. g., in urban landscapes not only to manage noise but also to actively design pleasant sound environments (PRIOR 2017, JIANG et al. 2018). Yet there are neither standard approaches how to integrate sound into landscape character assessments (PRIOR 2017) nor how to link sounds with authentic references to virtual landscape scenes (JIANG et al. 2018). In this paper, we present an approach how to reproduce environmental sounds for visual representations of different landscape types. We focus on the evaluation of these sound ambiences: 1) if people perceive them as consistent with what they are seeing and 2) if they support experiencing the demonstrated landscape.

There are different approaches for recording and reproducing sound ambiences. LINDQUIST et al. (2016), for example, paired real sounds recorded in a park with 3D landscape visualizations to evaluate empirically the effects of sound on realism and preference ratings of virtual landscapes. They show that aural and visual stimuli interact with each other and influence people's perception of landscape realism and their preference ratings. The more congruent sounds and visuals are, i. e., the more they reflect what people expect to perceive, the higher the realism and preference ratings. LINDQUIST et al. (2016) point out that in case of unknown specific locations – which would occur if generic landscape types are considered – people are responding to the visual contents of the simulations and their expectations of what to experience remembering situations in similar landscapes.

Instead of playing back the entire sound recorded in an environment, OLDONI et al. (2015) selected a subset of sounds that is most representative for a given location. They constructed an acoustic summary, which they validated in a listening test. One of their major findings is that soundmarks are very important in describing an acoustic environment. Soundmarks (in analogy to landmarks) are defined as sounds reflecting natural (e. g., waterfalls, natural wind traps) or cultural characteristics (e. g., distinctive bells, sounds of traditional activities), which are given particular attention by a community and are often the sounds first noticed (KANG 2007: 45, 98). Furthermore, selecting sounds by combining the saliency and the frequency of occurrence of sound events at a specific location resulted in acoustic summaries that are highly representative (OLDONI et al. 2015).

Starting from these insights for sounds representing specific locations, we developed an approach for collecting sounds and reproducing sound environments for landscape types. Further, we evaluated the perceived consistency of the sounds with the landscape types' visual representations in a laboratory experiment. Finally, we used these results to critically review the overall approach and to provide a starting point for coupling environmental sounds with visual simulations of landscape types.

2 Methods

2.1 Visual Landscape Representation

In total, we selected seven cultural landscape types based on the landscape typology of Switzerland (ARE, BAFU, BFS 2011; see Tab. 1). These landscape types are located in the different biogeographical regions of Switzerland, and they differ from each other in the settlement and land use types. Visual representations of these landscape types were prepared using panoramic photos from specific locations, which contain the characteristic elements of the respective landscape type but are not well known (Tab. 1). These 360-degree photos were taken at the seven locations using Google's app "Street View" on a smartphone.

Table 1: Characterization of the seven audio-visual stimuli (* = Soundmarks)

Landscape Type	Visualization	Sounds constituting the overall sound ambience	L_{Aeq}, 30 s [dB]
Jura		Flies*, cowbells*, bird of prey*, crickets chirping, soft wind, distant light traffic	38.50
Plateau urban		Slap of a car door*, church bell*, rooster*, rattles of a nearby flag, distant highway traffic, bird twitter, human steps	41.98
Plateau agriculture		Crows*, agricultural machinery*, church bell*, bird twitter, distant traffic	38.52
Prealps		Hay turners*, distant cow bells*, insects, distant airplanes, bird chirp, distant dog barking	36.44
Alps touristic		Flies*, marmot whistle*, hikers talking*, cableway*, bird chirp, soft wind, metallic rattles	42.11
Alpine valley		Heavy traffic*, train passing*, bird twitter, crickets chirping	56.87
Alps abandoned		Alpine bird chirp*, crickets chirping*, helicopter passing*, flies, creek rushing	42.60

2.2 Sound Recording and Processing

With a sound-field microphone (4 channel, first order ambisonics) we first recorded the environmental sounds at the exact point of view of the panoramic image of the different landscape types. We carried out recordings of about 20 minutes at different specific moments of the day (morning, midday, afternoon) while taking notes of the sound events. These descriptive recordings and listening protocols served as basis for identifying soundmarks. If the quality of the descriptive recording was not sufficient to isolate these specific sounds, we conducted another recording in a more isolated manner, e. g., recording the sounds of cowbells, of typical birds, of bicyclist, or of haymaking. Then, we selected, processed, and arranged the sounds for each landscape type utilizing the digital audio workstation REAPER (https://www.reaper.fm).

2.3 Experiment Design

Finally, we tested the audio-visual simulations to find out as how consistent people perceive the established ambient environmental sounds with the visualizations of the landscape types. Therefore, we designed a laboratory experiment. We presented the seven audio-visual stimuli to the participants for 30 seconds each in an AudioVisual Lab. The AudioVisual Lab is situated in a sound absorbing room and comprises a set of 20 loudspeakers (16 fixed on the walls and 4 on the ceiling for 2D and 3D sound reproduction settings). A surround sound processor (Sonic Emotion, www.sonicemotion.com) allows for placing virtual sound sources in the room controlling the directions of sound emission. In this case, the sounds were played back in a 5.1 surround setup. Implementing a sound level meter, we adjusted the sound pressure level to the value measured during recording in the field (Tab. 1). However, due to pre-testing results, we decided to reduce the volume by 5 dB for the landscape type "Alpine Valley" since the sound was perceived as too loud. The 3D visualizations were projected on three micro-perforated screens for a panoramic perception of the visual landscape scene (Fig. 1).

Fig. 1:
Setup of the experiment in the AudioVisual Lab: The participant in the middle of the room is looking at the panoramic projection of a landscape type and hearing an ambient sound prepared for this scene

Fifty-two people (34 females, 28 males) aged between 20 and 49 years (M = 29.51, SD = 7.65) and working or studying in different fields (spatial and landscape planning, environmental sciences, architecture, art, energy sector, administration) stated normal hearing and completed the experiment. Each participant perceived a set of the seven stimuli (Tab. 1). After each stimulus, the participant rated the perceived consistency of the presented situation on an eleven-point Likert Scale. Then, focusing only on the final stimulus presented, they answered further quantitative questions on the perceived realism, their immersion into the landscape scene, their perception of the sounds using bipolar adjectives (e. g., unpleasant – pleasant; loud – quiet; dominant – accompanying, etc.), and which sounds were most noticeable. Afterwards, we conducted a short interview asking qualitative questions to gain reasons for the participants' ratings. Finally, the participants answered socio-demographic questions and stated in which landscape type they spend most of their time, their leisure time, and in which landscape type they feel at home.

Overall, we defined seven participant groups, whereby within a group the last stimulus was held constant. The other six stimuli were randomized over all groups implementing the Latin Square technique in order to minimize sequential distorting effects. We calculated the randomization with the software R using the «crossdes» package's function «des.MOLS» (https://cran.r-project.org/web/packages/crossdes/crossdes.pdf).

The experiment was conducted with each participant individually, which took about 25-30 minutes.

2.4 Data Analysis

The software SPSS 24 was employed for analysing the quantitative data. Further, we transcribed the interviews. Then we coded the content with regard to the following questions: Which sounds did the participants notice first, respectively, which sounds stand out and why? What did the participants associate with the perceived simulation? Why did which sounds support or prevent the landscape becoming alive? What was decisive for the consistency rating? This content analysis resulted in a summary of the answers.

3 Results

Participants rated the representation of the landscape type "Jura" (M = 9.98, SD = 1.394) significantly higher for the consistency of sound with visualization than all other landscape types' representations (Fig. 2). The types "Plateau urban" (M = 6.79, SD = 2.154), "Alps touristic" (M = 6.92, SD = 2.213), and "Alpine valley" (M = 6.48, SD = 2.297) show the lowest median values, and for the latter landscape type the values have the widest dispersion. The consistency ratings of "Plateau agriculture" (M = 7.69, SD = 1.853) and "Alpine abandoned" (M = 7.35, SD = 2.057) have quite similar mean values to the overall mean of the ratings (M = 7.49). The difference in the consistency ratings between the landscape types was tested with a one-way ANOVA for a 95 % significance level. The consistency rating for "Jura" differs significantly from "Plateau urban" ($p = < 0.000$), "Alpine touristic" ($p = < 0.000$), "Alpine valley" ($p = < 0.000$), and "Alpine abandoned" ($p = .007$). In addition, the consistency rating for "Prealps" differs significantly from the rating of "Plateau urban" ($p = .037$) and "Alpine valley" ($p = .033$).

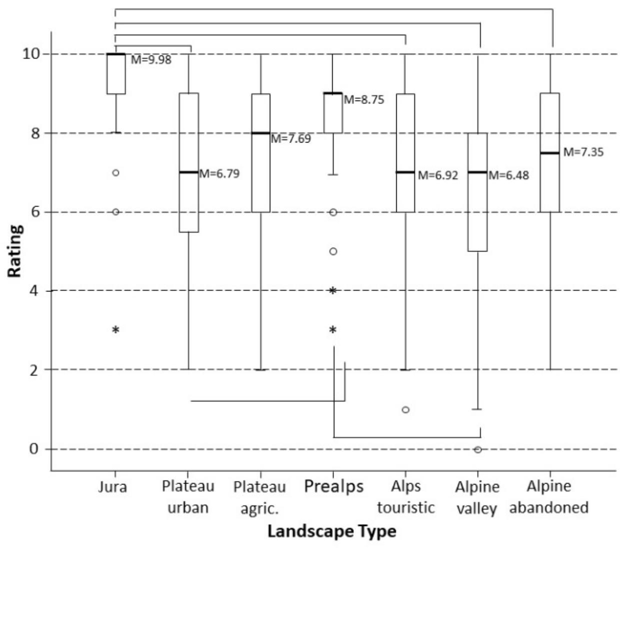

Fig. 2:
Rating of the consistency of the audio-visual stimuli of the seven landscape types

The rating of the perceived level of realism of the sounds as well as the rating of the participants' perceived immersion into the landscape scene show the tendency of being positively correlated with the overall perceived consistency of the audio-visual stimuli. However, the correlation is not significant.

The ratings of the bipolar adjectives show that the sounds of "Alpine Valley" were perceived very differently from the sounds of all other landscape types (Fig. 3). They were perceived as rather unpleasant, loud, and dominant. In contrast, sounds for "Jura" were rated as very pleasant, accompanying, real, and appropriate. Further non-parametric Spearman correlation analysis of the bipolar adjectives' ratings demonstrates a significant correlation between the consistency rating and the sounds' perceived volume ($p = 0.001$) and dominance ($p = -0.955$). The analysis of the participants' statements in the interviews provided insights into their perception of the sounds and reasons for the ratings.

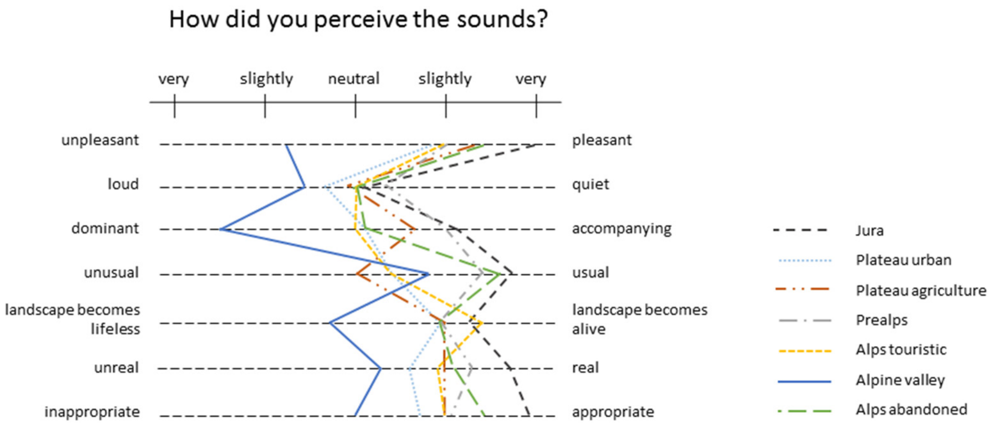

Fig. 3: Rating of the bipolar adjectives for the representations of the seven landscape types

Most participants named one of the soundmarks as sounds they noticed first or as salient sounds. They identified the landscape types correctly or described them as ordinary situations in Switzerland. According to the participants' statements, the soundmarks substantially enhanced the authenticity and vividness of the landscape. However, it was similarly important to them to perceive the overall sound ambience, i. e., all sounds together, as harmonic and consistent with the landscape image. This was achieved for the "Jura", whereas for all other landscape types inconsistencies were mentioned.

Decisive for the consistency rating was primarily the matching of what they were seeing and hearing. Consistency depends on the possibility to locate the source of the sounds, on the suitability of the sounds regarding the time of the year and day, and on whether the sounds fit to the visual dynamic in the images. For example, the still images irritated a couple of participants because they heard rustling leaves but they did not see them move. Equally important for the rating was which sounds the participants anticipated with regard to the shown situation. They said that they compared the sounds heard with the ones they expected. Sounds they did not expect, e. g., the rooster in "Plateau urban", or that they could not clearly identify, e. g., the creek rushing in "Alps abandoned", affected their perceived consistency negatively. Also sounds they associated with dangerous or uncomfortable situations, e. g., the helicopter in “Alps abandoned” or the flies around the head in “Alps touristic”, were perceived as unpleasant.

Further, many participants said that the volume and dominance of the sounds were significant characteristics, which needed to be plausible regarding the visible situation. Hence, not only the obvious sounds need to be audible but also other, more general ambient sounds, such as the hikers talking in “Alps touristic”. Some participants stated that this made the simulation more realistic. For “Alpine valley”, the volume and dominance were perceived as not at all consistent, because most participants thought that the volume of the traffic noise was too high and dominated the sound ambience due to the permanent, monotonous sound without interruption.

4 Discussion

Overall, the ratings show that the sound ambiences were already quite consistent for the landscape types “Jura” and “Prealps”. These landscape types have a great visual similarity (mosaic of pastures and forest), and the sounds constituting the sound ambience are also quite similar (cowbells, bird chirp, insects/flies). Participants perceived these sounds as usual, real, and appropriate. However, some participants stated also for “Alpine valley”, “Alps touristic”, and “Alps abandoned” that they could close their eyes and imagine being in such a landscape. Since all participants mentioned the soundmarks and as these are characteristic for the landscape types, they seem to support associations with the landscapes presented in the simulations. However, also the other ambient sounds which are more in the background and not landscape type specific seem to be equally essential to present a realistic overall sound ambience. Because expected sounds were perceived as “normal” and therefore consistent, rather ordinary sounds should be integrated into this mix. However, this might result in a trivialisation of the sound ambience, which is not adequately reflecting a landscape type’s character (PRIOR 2017).

Based on the participants’ statements, in pleasant, consistent sound ambiences the sounds blend into each other so that no sound is dominant with regard to its frequency and its volume, and the whole ambience accompanies the visual landscape. This means that according to the distance to the landscape elements the respective sounds need to be audible in an intensity perceived as appropriate. This is a critical insight, because the sounds’ volume was adjusted to the one measured in the field. Only for “Alpine valley” the volume was already decreased by 5 dB as the sound was perceived as too loud in the pre-test (see section 2.3). Apparently, the sounds’ volume is not equally perceived being in the landscape or being in a laboratory. As this was particularly true for “Alpine valley”, which was characterized by heavy traffic, it is not clear how far the perceived “pleasantness” may also play a role in this context (see also AXELSSON et al. 2010).

With regard to the evaluation method, we found the experiment design combining quantitative and qualitative information for assessing the consistency of the sound with the landscape visualizations very effective. In the interviews in particular, the participants provided helpful insights on what gathered their attention and why they perceived sounds as consistent or not. Furthermore, the duration of the experiment was said to be adequate. In order to keep the participants focused, the duration should not be extended significantly. However, according to the participants' feedback it was not always clear to them what to expect in the beginning. Therefore, a test-example of the audio-visual stimuli and how it is rated should be given. Furthermore, the participants should be asked, which sounds they expected, in order to get clues, which sounds might be missing in the sound ambience. To test, calibrate, and validate the participants' responses, stimuli could be used where visuals and audio do not match. In this way, e. g., it can be tested, how well people remember the sounds they perceived in similar situations of the presented landscape types.

5 Conclusion and Outlook

The goal of this study was to combine environmental sounds with visual representations of different landscape types and to test how well they are perceived as consistent and support experiencing the demonstrated landscape. Our major findings confirm LINDQUIST et al. (2016) and show that it is important to provide sounds that are expected and whose sources can be located in the visual landscape representation. Soundmarks seem to support people's association with the shown landscape type, but it is the right mix of a consistent frequency of soundmarks with ambient sounds in the background that fosters consistency. This goes in line with the findings of OLDONI et al. (2016). Furthermore, the sounds' volume and dominance are crucial factors influencing the perceived consistency. Thereby, the perception of the sound volume seems to be different in the laboratory situation as opposed to the situation in the field.

Further research in this direction is urgent, as recent perception studies of landscape changes require practical guidance for valid audio-visual simulation. This is particularly relevant considering that long-term, far-reaching policy decisions are based on such study outcomes. In a next step, the visual simulation will be based on point cloud data in order to enhance the generic representation of the landscape types. Due to the pointillist visualization style, an abstraction of the specific location is achieved, while still providing a rather photorealistic landscape rendering required for preference studies (SPIELHOFER et al. 2017). Furthermore, for even more enhancing the immersion into the landscape scene, virtual reality glasses can be employed (WISSEN HAYEK et al. 2016). It should be tested, which effect the different presentation modes (panoramic presentation vs. virtual reality glasses) have on the perceived consistency of the audio-visual stimuli as well as finally on the preference ratings of landscape changes.

Acknowledgements

This study was carried out in the scope of the NRP 70 project ENERGYSCAPE (Research Grant: 407040_173808 / 1).

References

ARE, BAFU, BFS (2011), Landschaftstypologie Schweiz. https://www.are.admin.ch/are/de/home/laendliche-raeume-und-berggebiete/grundlagen-und-daten/landschaftstypologie-schweiz.html.

AXELSSON, Ö., NILSSON, M. E. & BERGLUND, B. (2010), A Principal Components Model of Soundscape Perception. The Journal of the Acoustical Society of America, 128 (5), 2836-2346.

CELIO, E., OTT, M., SIREN, E. & GRÊT-REGAMEY, A. (2015), A Prototypical Tool for Normative Landscape Scenario Development and the Analysis of Actors' Policy Preferences. Landscape and Urban Planning, 137, 40-53.

KANG, J. (2007), Urban Sound Environment. Taylor & Francis, London/New York, 278.

JIANG, L., MASULLO, M., MAFFEI, L., MENG, F. & VORLÄNDER, M. (2018), A Demonstrator Tool of Web-Based Virtual Reality for Participatory Evaluation of Urban Sound Environment. Landscape and Urban Planning, 170, 276-282.

LINDQUIST, M., LANGE, E. & KANG, J. (2016), From 3D Landscape Visualization to Environmental Simulation: The Contribution of Sound to the Perception of Virtual Environments. Landscape and Urban Planning, 148, 216-231.

OLDONI, D., DE COENSEL, B., BOCKSTAEL, A., BOES, M., DE BAETS, B. & BOTTELDOOREN, D. (2015), The Acoustic Summary as a Tool For Representing Urban Sound Environments. Landscape and Urban Planning, 144, 34-48.

PRIOR, J. (2017), Sonic environmental aesthetics and landscape research. Landscape Research, 42 (1), 6-17.

RIBE, R. G., MANYOKY, M., WISSEN HAYEK, U., PIEREN, R., HEUTSCHI, K. & GRÊT-REGAMEY, A. (2017), Dissecting Perceptions of Wind Energy Projects: A Laboratory Experiment Using High-quality Audio-visual Simulations to Analyze Experiential Versus Acceptability Ratings and Information Effects. Landscape and Urban Planning, 169, 131-147.

VAN BERKEL, D. B. & VERBURG, P. H. (2014), Spatial Quantification and Valuation of Cultural Ecosystem Services in An Agricultural Landscape. Ecological Indicators, 37 A, 163-174.

VAN EETVELDE, V. & ANTROP, M. (2009), A Stepwise Muti-Scaled Landscape Typology and Characterization for Trans-Regional Integration, Applied on the Federal State of Belgium. Landscape and Urban Planning, 91, 160-170.

VAN DER ZANDEN, E. H., LEVERS, C., VERBURG, P. H. & KUEMMERLE, T. (2016), Representing Composition, Spatial Structure and Management Intensity of European Agricultural Landsapes: A New Typology. Landscape and Urban Planning, 150, 36-49.

SPIELHOFER, R., FABRIKANT, S. I., REBSAMEN, J., VOLLMER, M., GRÊT-REGAMEY, A. & WISSEN HAYEK, U. (2017), 3D Point Clouds for Representing Landscape Change. Journal of Digital Landscape Architecture, 2-2017. Wichmann, Berlin/Offenbach, 206-213.

SWANWICK, C. (2002), Landscape Character Assessment. Guidance for England and Scotland. The Countryside Agency, Scottish Natural Heritage. http://www.snh.org.uk/pdfs/publications/LCA/LCA.pdf.

WISSEN HAYEK, U., WALTISBERG, D., PHILIPP, N. & GRÊT-REGAMEY, A. (2016), Exploring Issues of Immersive Virtual Landscapes for Participatory Spatial Planning Support. Journal of Digital Landscape Architecture, 1-2016, Wichmann, Berlin/Offenbach, 100-108.