



COMPUTATION:
THE NEW REALM OF ARCHITECTURAL DESIGN

eCAADe27
istanbul 2009

edited by Gülen Çağdaş and Birgül Çolakoğlu

Using Audio-Augmented Reality to Assess the Role of Soundscape in Environmental Perception

An Experimental Case Study on the UC Berkeley Campus

Gökçe Kınayoğlu

College of Environmental Design, University of California at Berkeley, USA

gokce@berkeley.edu

Abstract: *Sounds, along with other senses, have a profound influence on our perception of the environment. The multi-modality of perceptual processing is influential in cognitive interpretation, semantic and aesthetic evaluations of environmental scenes. This paper describes an experimental case study using audio-augmented reality, carried out in order to better understand how sound influences sense of place. A correlation is established between soundscape and sense of place that depends on audio-visual congruence based on cultural, aesthetic and semantic factors. Subjective influences of soundscape on place experience are grouped and discussed under 5 areas that were possible to identify: Emotive and synaesthetic effects; effects on attention, gaze and behavior; effects on spatial orientation and sense of scale; influence of audio-visual congruence on sense of place; and perception of personal and social space.*

Keywords: *Augmented reality; soundscape; environmental perception; place theory; environmental acoustics.*

Introduction

Phenomenological accounts of perception suggest that intentionality and embodiment are integral to the sensory experience of the world (Merleau-Ponty 2002 [1945]). Similarly, Heidegger argues that human experience is place-centric, with his concept of *Das-sein*, or “being-in-the-world” (Heidegger 1996 [1953]). In line with the arguments of phenomenology and existential philosophy, investigations on the psychology of perception also suggest the importance of the environmental context and sensorimotor

interaction in perceptual processes (Gibson 1979). Accordingly, the Cartesian paradigm which formulates perception as passive interpretation of a stream of sensory information flowing inwards from the world through our senses and into our minds gets discredited in favor of a mobile-intentional subject model that actively explores and probes the environment in constant interaction with it (Noe, 2004).

Augmented reality simulation techniques resonate well with the ecological paradigm of perception. Unlike fully-mediated virtual environments, which operate as complete substitutes masking out

and replacing immediate experiences with synthesized ones, augmented reality works by combining physical experience of space with digital interactivity. Augmented reality is defined with three characteristics: 1) Combines real and virtual content. 2) Allows for real-time interactivity. 3) The virtual content is associated to and registered with the real world (Azuma et al., 2001). By coupling the realism of the physical setting with the versatility of digital media, augmented reality reincorporates body and its interactions with physical space into computational representation.

Existing research on audio-visual interaction in environmental settings

Existing studies suggest that perceptual processing of auditory and visual information are interconnected. The multi-modality of perceptual processing is

influential in cognitive interpretation, semantic and aesthetic evaluations of environmental scenes.

As an example, a strong correlation between the type of visual scene and subjective evaluations of loudness has been demonstrated. It has been shown through an experiment using sound-image pairs that red trains are perceived to be louder than blue ones (Patsouras et al., 2002); sound of white noise is reported to feel more pleasant when simultaneously presented with a picture of a waterfall (Abe et al., 1999); images of greenery reduce the perceived loudness of noisy soundscapes (Fastl, 2004). A large scale survey made with a group of people living in the vicinity of wind turbines showed that the amount of noise annoyance is positively correlated to the visual presence of the turbines (Pedersen and Larsman, 2008).

Using image-sound pairs, Viollon et al. (2002) examined the effect of the level of urbanization on the

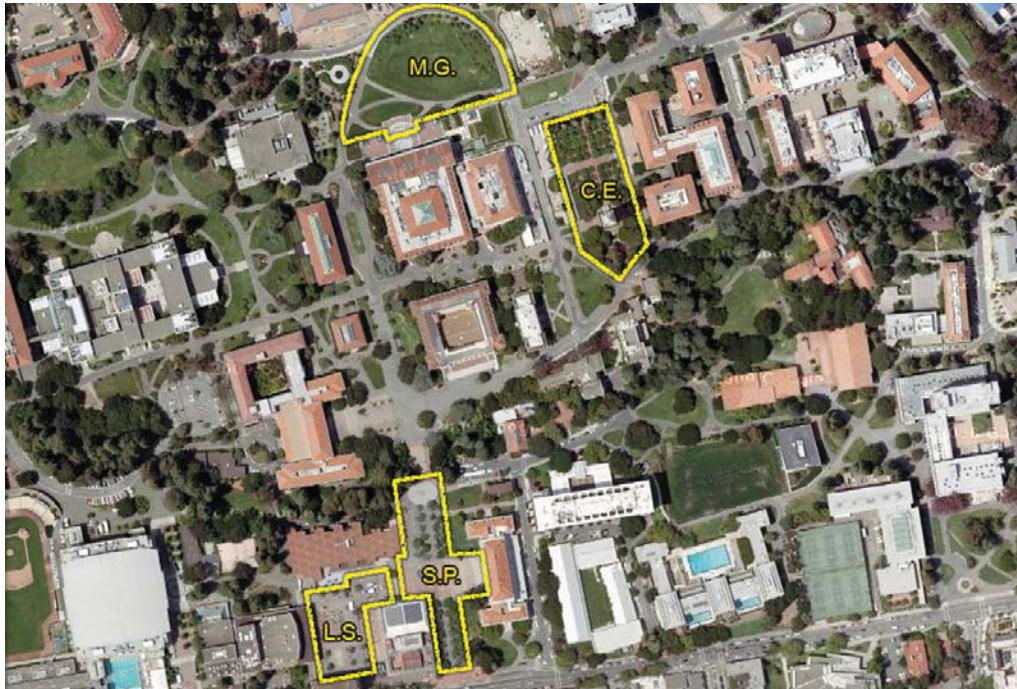


Figure 1
Study locations at UC Berkeley Campus: Upper Sproul Plaza (S.P), Lower Sproul Plaza (L.S.), Campanile Esplanade (C.E.), Memorial Glade (M.G.) (Source: Google Maps)

perception of various kinds of noise. They showed that the more urbanized the visual environment, the less tolerable the same amount of noise became.

Carles et al. (1999) carried out an experiment with sound-image pairs in order to determine how different sounds influence the aesthetical perception of visual landscapes. The results indicated that: 1) Coherent combinations of sound-image pairs were more likely to get higher ratings. 2) Natural sounds improved the ratings of the images regardless of the coherence within the image-sound pairs. 3) Aesthetically unpleasant sounds such as traffic noise negatively influenced the overall ratings of sound-image pairs (Carles et al., 1999).

Methodology and implementation

This study was carried out as a series of surveys conducted at the UC Berkeley campus with a group of 12 participants, all graduate students from the College of Environmental Design. The surveys took place at four different locations on campus (Figure 1). Each participant was randomly assigned to visit three locations on campus. For two out of the three places, the participants were asked to wear a mobile audio-augmented reality device and freely explore their surroundings by walking. The third location was explored in an unmediated manner. After visiting each environment, participants responded to a one-page survey and a short interview.

The mobile audio-augmented reality system was designed to deliver its wearer a geo-located, immersive acoustic experience. The participants were able to see, smell and touch the physical environment and freely move around while wearing the headphones connected to the system. By dynamically tracking the users' geo-location and head direction, the system presented a stable virtual soundscape with simulated spatial and directional cues.

The study examined the effects of the soundscape on environmental perception in situations where visual and acoustic information were in varying degrees of congruence with each other. The

classification of combinations was based on the aesthetic characteristics of the sounds themselves and their semantic and cultural compatibilities with the existing places.

Real-time outdoor spatialization of augmented-audio

Audio-augmented reality was initially envisioned in 1993 (Cohen et al., 1993). Many implementations and usage areas for augmented audio have been suggested and/or developed such as an audio-guide in an indoor art-museum context, free-field applications for non-visual low-attention navigational interfaces, or for geo-located sound art.

The mobile audio-augmented reality setup that was developed for this experiment ran on a notebook computer equipped with a GPS receiver (GlobalSat BU-353 USB) for motion tracking, headphones (Sony MDR7505), and a 3-axis digital compass and a tilt sensor for head tracking (OceanServer OS5000-US). Soundscape authoring and spatial audio simulation via psycho-acoustic signal processing was carried out by a custom patch built on version 5.0 of the Cycling74 MAX/MSP signal-processing environment.

Numerical ratings

The soundscape treatments for each location were classified according to their intrinsic acoustic pleasantness and their compatibility with the visual environment in semantic and cultural terms. Influence of each of these treatments on participants' environmental perception and preferences were queried for factors of pleasantness, vibrancy/impressiveness, noisiness, relaxation, orientation, familiarity, personal/intimateness, and familiarity of the environment.

Upper Sproul Plaza: The intent of this part of the experiment was to test how culturally alien soundscapes can influence sense of place. The sounds chosen to be overlaid on top of Upper Sproul Plaza were sounds of a traditional Moroccan marketplace, with sounds of the market crowd with



Figure 2
 (left, middle) Participants using the augmented audio reality system
 (right) map showing locations and audible ranges of the virtual sound sources

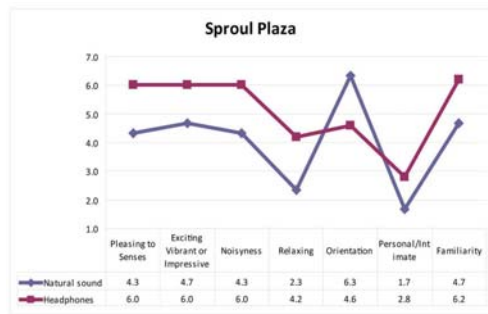
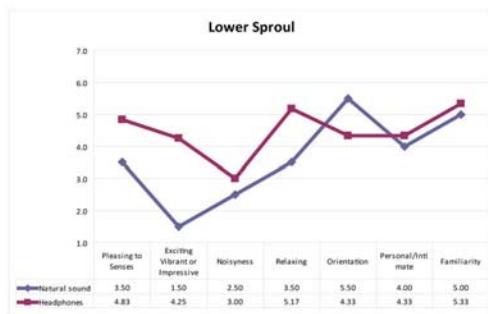


Figure 3a, 3b
 Images and survey results from Sproul and Lower Sproul Plazas



audible conversations in Arabic, ethnic music playing, sounds of metal smiths and craftsmen, and cattle. The expected outcome was acoustic alienation, a decreased sense of familiarity. A change in pleasantness, excitement or impressiveness was not expected since the overlaid sounds were aesthetically neutral but solely chosen to be “out of place.”

The majority of the participants did not report a big discrepancy between the sounds and the visual landscape. On the contrary, substantial increases in the aesthetic preference ratings and some increase in familiarity were recorded (Figure 3a). The congruence between the semantic characteristics of the audio and visual inputs and the sensory awareness of human activity resulted in an increase in the pleasantness of the overall audio/visual ambiance.

Another interesting outcome was that despite the fact that the plaza was perceived to be noisier with the headphones on, it was also reported to be more relaxing. This may be attributed to the use of headphones causing a feeling of isolation from the environment making one feel insulated in his/her private “bubble”, an effect which is also observed in personal stereo users (Bull, 2000). Correspondingly, people found the environment more personal/intimate with the headphones on, than without.

Lower Sproul Plaza: Sounds that were introduced through the augmented audio system intended to create a naturally pleasing sonic atmosphere were overlaid on the urban visual characteristics of the Lower Sproul. Looping ocean wave sounds were placed in a row along the west side of the plaza. Sounds from a forest, consisting of birdsongs and wind in leaves, were placed on the east. The augmented soundscape was intended to be aesthetically pleasing, culturally neutral yet semantically incompatible with the visual environment.

The results indicate that the strongest influence was on the aesthetic attributes and comfort values (Figure 3b). This finding was in line with the expectations. The mean values for responses to questions regarding aesthetic preferences had the most profound shifts, in “pleasantness” and “exciting/vibrant/

impressive” ratings. This shows that the soundscape elements, although they were not visually in tune with the environment, generated positive influence on the overall character of the environment.

Campanile Esplanade: This environment served as a test case for which the acoustic treatment was designed to match the visual landscape as closely as possible, only enhancing or making use of the sonic affordances existing on site. For instance, carillon music originally recorded at the Sather Tower (Campanile) was played back from the virtual location of the tower. Nature sounds, sound of trees, chirping birds and crickets, were scattered around the park.

Mean ratings for the groups with and without headphones closely overlapped. The acoustic treatment in this case caused no significant change (Figure 4a).

Memorial Glade: The audio environment overlaid on top of Memorial Glade was that of a crowded urban plaza. Sounds of traffic were placed at the north and western edges of the lawn. Sounds of a reverberant bus station interior with announcements and crowd noises overlapped the Doe Library steps. Loud noises from a construction site were over the northeast quadrant of the plaza. The sounds of the Sather Tower carillon were designed to be audible at the southeast corner of the site where it borders the nearby Campanile Esplanade.

The overall impact of the new sonic ambiance was aesthetically and semantically negative. Nearly all participants reported the environment to be much less pleasing, much less exciting or impressive, even though the sonic activity level was higher than what is actually on site (Figure 4.b). Noisiness was reported to increase and the environment was reported to be much less relaxing than it is with the actual soundscape. It was also perceived as much less personal/intimate, and more public, and less familiar.

Subjective feedback

In the discussions that follow participants are identified with their numbers from P1 to P12. Location

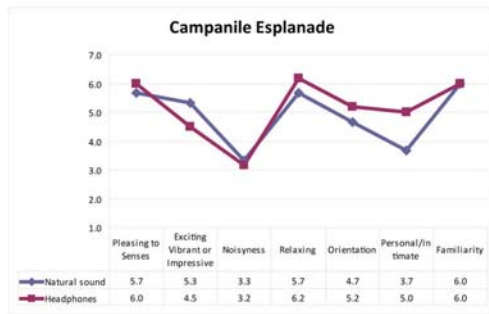


Figure 4a, 4b
Images and survey results
from Campanile Esplanade
and Memorial Glade

names are abbreviated as follows: Campanile Esplanade (C.E.), Memorial Glade (M.G.), Upper Sproul Plaza (S.P.), and Lower Sproul Plaza (L.S.).

Emotive and synaesthetic effects

Sound can transform the emotion or mood associated with a place. In some cases perception of sound can also trick the other senses and induce synaesthetic effects:

"The birds singing made me feel as if it was a sunny day, yet today is a very cloudy day... The Campanile [Esplanade] feels like a sunny day, like Alhambra Gardens." (P2, C.E.)

"The water sounds made it seem physically cooler." (P9, C.E.)

At Memorial Glade, where the overlaid soundscape was that of an urban street environment sounds of traffic and construction put the participants on alert and made them look out for vehicles

that may be around. Some participants experienced a feeling of uneasiness because of this. Others reported that they literally became concerned for their personal safety:

"The traffic sounded more constant, steady. Sounded like I am somewhere in the middle of the street. I did not like hearing traffic sounds. They made me have to be real careful." (P6, M.G.)

An opposite effect was observed in Lower Sproul Plaza. In contrast with the traffic sounds played at Memorial Glade, the sonic ambiance of a seaside superimposed with the visual experience of the Lower Sproul had a physically comforting effect. One participant described the sounds as being "depressive, in a positive way":

"... they lower your heart rate down... The ocean sounds are very soothing... Although Lower Sproul is not an active space – it can feel quite isolating. Yet I felt safe. The water crashing on the shore is a

safe sound for me... Also this type of sound – and therefore the space – is intimate in the sense that it encourages introspection of the self.” (P11, L.S.)

Effects on attention, gaze and movement

Whether they may be perceived as fully realistic or overtly artificial, foreground sounds made the participants attend to their perceived locations in search for their potential sources. The participants experimented with the virtual sound environment by turning their gaze and bodily orientation towards the sounds and walking around the space to move towards locations where they thought the sounds may be originating from:

“The sound of the Campanile made me want to look up. Even though I knew that it was not really playing I looked up to see the source of the sound.” (P2, C.E.)

“Even it’s a well designed park, very symmetric and directed, I feel like the wild. I would like to sit and be quiet. The presence of the campanile and the bell sounds made me look up to the sky. Feels good!” (P4, C.E.)

The sounds were designed to be in overall harmony with the environment on the Esplanade yet it is clear that they highlighted certain aspects of the place and suppressed others. Sounds directed the participants’ attention, gaze and movements towards particular details, objects and locations. Certain features of the environment that would otherwise go unnoticed came into the attention of the participants while they were looking for the sources of the sounds.

Effects on spatial orientation and sense of scale

Sound localization worked best when the artificial acoustic sources were matched with existing visual objects. Virtual carillon music coming from Sather Tower (Campanile) was an excellent illustration of this phenomenon:

“I could not tell whether the Campanile was really ringing. I thought it was from the speakers but

I could not tell, really. It was hard, until I sort of removed [the headphones] to check. The directionality was really nice, that is why I was surprised.” (P5, M.G.)

“The Campanile was REALLY here. It felt more strongly spatialized.” (P9, C.E., original emphasis)

Precise sense of localization like this was not attained when there are no visual cues to anchor sounds in place. This was mainly due to two reasons. First, because of the technical limitations of the location-tracking system it was only possible to achieve a low accuracy representation of source azimuth. Second, the absence of visual sources to match the sounds made them ecologically invalid, or “ghost” sounds. Even so, the study participants were able to attain an overall feel of how the sounds are distributed in relation to each other, and construct a mental map of their locations:

“Sounds stayed at the same location when I rotated my head. This was very useful for orientation.” (P10)

“I was able to map their sources in space because sounds changed as I move.” (P9)

“I found myself trying to localize the ironsmith shops and the spatial configuration of the tables at Sproul help associate some kind of SCALE to the sounds. (P9, S.P.; original emphasis)”

Sounds without a visually specific point of origin were registered as ambient acoustic zones. These sounds were interpreted as spaces that were experienced from within, rather than objects or events perceived at a distance.

Audio-visual congruence and sense of place

In this study a complete representational realism was not an expected outcome. Visual information and audio was intentionally out of tune and out of sync with each other. In three out of the four sites, the artificially generated soundscape was not semantically compatible with the visual landscape. Nevertheless the participants were asked about their feelings on how well the sounds blended with the visual atmosphere. They were also asked to provide feedback on effectiveness of the interface:

"The sounds were natural, but the floor was concrete. Disassociation of visuals and audio did not make sense. In Sproul, this made sense because there were people around and I heard people." (P10, L.S.)

"I was totally in one place... It was like really compelling, dissolved in one..." (P4, C.E.)

"I could not tell whether it was real or not, when the sounds matched the visuals." (P6)

When the participants heard noises of things or events that were not visually present, they tended to construct valid explanations to make sense of their environments. Interestingly, the effect was not completely detrimental to the sense of place. Like off-screen sound in a film, invisible sound sources were either thought to be coming from the space behind the observer, or to be occluded by objects, buildings or landscaping in between:

"The sounds made it feel quite busy. Feel like there must be a lot going on behind the trees that border the glade." (P5, M.G.)

"But then when I was here, in Sproul, there were all these diverse [sounds] outside, for sure, non-American sounds it made me feel that it was not there. So it was not that compelling... But then I thought maybe it is in my back." (P4, S.P.)

Effects on perception of personal and social space

The perception of sonic objects and events that are not emerging from the visual atmosphere influenced the overall sense of place reported by the participants. Not being able to hear the sounds of the actual environment also influenced how the environment is experienced. The participants reported a decrease in their connectedness with the environment:

"You usually hear things in the environment like footsteps and other details. They disappear with the headphones... The people seemed to be on 'mute'." (P9, C.E.)

Being detached from the actual acoustic atmosphere of a place induced a feeling of personal

isolation and spatial alienation. However this was not always commented upon as a negative phenomenon. Some people liked the sense of acoustic disconnection, and others mentioned that they preferred listening through the headphones to the actual soundscape of the places.

"Very interesting experience. I felt more relaxed and felt that Sproul was intimate! The effect of the sounds in my ears created a bubble for me – I felt detached from my surroundings, in a positive way. A similar experience to when I run with my iPod. (P11, S.P.) The sounds with the headphones felt like actual Sproul. When I took the headphones off it felt too silent, inactive." (P1, S.P.)

"The place feels mundane, after taking off the headphones." (P6, L.S.)

People who reported that they regularly use personal stereos experienced less trouble adjusting to the artificial soundscape compared to people who do not wear headphones or earphones on a daily basis. Relative ease with using headphones may be explained either as the result of a learned behavior due to previous habituation with the use of personal audio devices or as a general preference of whether the participant likes or dislikes using them.

Conclusion

This study demonstrated that soundscapes do influence characteristics of place experience. Synaesthetic influences of sound on other senses were shown to alter the experience of the same environment to feel physically different, colder, windier, chillier or sunnier and warmer than it is. An example of the emotive effects of soundscape is the influence on the stress level. Sounds were shown to influence attention, behavior and movement within the environmental context. Gaze, head and bodily motions were all visibly correlated to elements of the augmented soundscape. Acoustic perception was shown to affect the sense of geometric, social and personal space.

The preliminary classification of the augmented

soundscapes included in this experiment in terms of their aesthetic, semantic and cultural characteristics were later correlated to the type of effect they induced upon the sense of place. Aesthetic qualities associated with the sounds were the strongest factors which in turn influenced emotional responses towards the environment. Semantic compatibility was found to be the most influential for judgments related to audio-visual congruence. When their informational content was not in tune with the visual environment, sounds became distracting. Audio-visual incongruence resulted in a level of anxiety and sense of disconnectedness even when the sounds were aesthetically pleasing by themselves.

All four locations included in this study were outdoor environments. Even though it would have been interesting to include indoor spaces in the study, this was not possible due to the reliance of the GPS on satellite connection. Future implementations of the study can be carried out with more sophisticated and accurate forms of position tracking, such as differential GPS for better outdoor localization accuracy, or an electromagnetic system that would enable indoor use.

Acknowledgements

This project was carried out as part of my doctoral dissertation directed by Dr. Yehuda Kalay at UC Berkeley. Special thanks to Dr. David Wessel and Adrian Freed from Center for New Music and Audio Technologies (CNMAT) for their help with the development of the hardware interface and audio spatialization algorithms.

References

Abe, K., Ozawa, K., Suzuki, Y. and Sone, T.: 1999, The effects of visual information on the impression of environmental sounds, *Inter-noise 99*, pp. 1177–82.
 Azuma, R., et al.: 2001, Recent advances in augmented reality, *Computer Graphics and Applications*, IEEE 21(6), pp. 34-47.

Bull, M.: 2000, *Sounding Out the City: Personal Stereos and the Management of Everyday Life*, Berg, Oxford ; New York.
 Carles, J. L., Barriola, I. L. and Lucio, J. V. D.: 1999, Sound influence on landscape values, *Landscape and Urban Planning* 43, pp. 191-200.
 Cohen, M., Aoki, S., and Koizumi, N.: 1993, Augmented audio reality: Telepresence/AR hybrid acoustic environments, in *IEEE International Workshop on Robot and Human Communication*.
 Fastl H.: 2004, Audio-visual interactions in loudness evaluation. in *ICA, Japan, Tu3.X1*.
 Gibson, J.: 1979, *The Ecological Approach to Visual Perception*, Lawrence Erlbaum, New Jersey.
 Heidegger, M.: 1996 [1953], *Being and Time: A Translation of Sein und Zeit*, trans. Joan Stambaugh, State University of New York Press, Albany.
 Merleau-Ponty, M.: 2002 [1945], *Phenomenology of Perception*, Routledge, London.
 Noe, A.: 2004, *Action in Perception*, MIT Press, Cambridge Mass.
 Patsouras, C., Filippou, T. G. and Fastl H.: 2002, Influences of color on the loudness judgment, *Forum Acusticum*. PSY-05-002-IP.
 Pedersen, E. and Larsman, P.: 2008, The impact of visual factors on noise annoyance among people living in the vicinity of wind turbines, *Journal of Environmental Psychology* 28(4), pp. 379-389.
 Viollon, S., Lavandier, C. and Drake, C.: 2002, Influence of visual setting on sound ratings in an urban environment, *Applied Acoustics*, 63(5), pp. 493-511.