Promenade

A Minimalistic Sculptural Interface for Immersive Soundscape Presentation

Martin Rumori
Institute of Electronic Music and Acoustics
University of Music and Performing Arts
Graz, Austria
Email: rumori@iem.at

***Abstract*—*Promenade* proposes an interface for the semi-personal, interactive presentation of soundscapes that allows for the seamless integration of visual material. It combines the excitation of a flat surface using transducers with transaural cross-talk cancelled stereophonic projection. The unfolding of a wide spatial image within the sweet spot is contrasted by a collapsed, localised impression outside. This inherent interactive behaviour may be further extended by means of a distance sensor.**

***Promenade* emerged from an experimental, arts-based research project. It therefore targets contexts that feature custom audio content rather than aiming at general-purpose applications requiring optimal acoustic properties.**

I. Introduction

Exhibitions or other presentations of distinct objects often call for an unobtrusive sound projection means that affords the visitors an immersive, mostly individual listening experience. Achieving such an experience requires spatial capabilities of the chosen projection technique.

In many such cases, headphones are used. Spatially rich sound material may be projected in particular with binaural signals. A major side effect is the inherent reconfiguration of our multi-modal perception, as the surrounding auditory domain is largely replaced by the projected one. Additionally, headphones may be considered as cumbersome or unhygienic in public exhibitions. Multichannel loudspeaker setups, in contrast, do not require physical engagement with the projection medium. However, they do not facilitate individual listening situations and they have to be mounted in an often complex constellation around the target listening area.

Promenade is a minimal auditory display prototype that relates to such issues. It is important to notice that *Promenade* has not been developed to overcome the limitations of other projection means. Instead, it emerged from a series of experiments in an arts-based research context as part of a sound installation with a focus on aesthetic listening experience. Due to its simplicity, its purpose and the idiosyncratic technology it incorporates, *Promenade* is not a universal, general-purpose spatial sound projection system but it manifests certain limitations. However, it might provide an appropriate and elegant solution in some of the circumstances described above. *Promenade's* qualities with respect to those circumstances are regarded retrospectively here in order to explore further possible fields of application and obvious modifications for different forms of appearance.

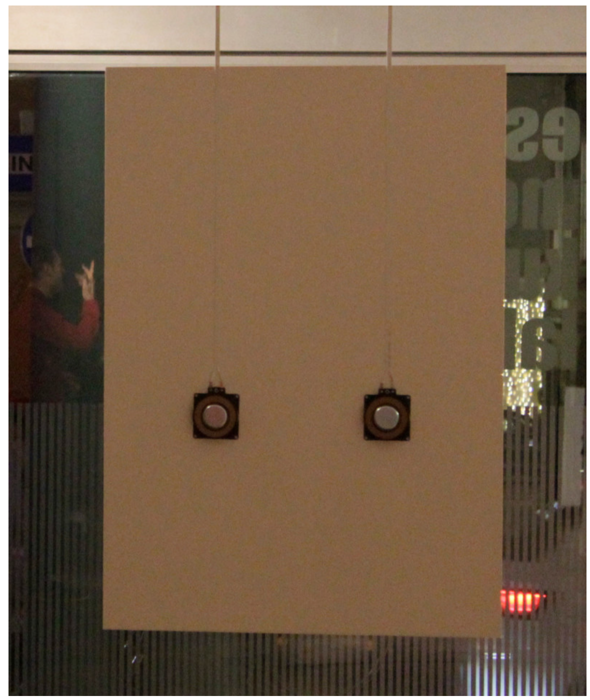

Figure 1. *Promenade* sound projection device.

This paper describes the basic idea of *Promenade* and its originating context (section II), followed by introducing the two main techniques involved, transaural crosstalk-cancelled sound projection (section III) and the principle of flat panel loudspeakers (section IV). Finally, some findings that arose from practical experience are discussed, along with possible modifications and extensions (section V).

II. Promenade

The *Promenade* sound projection device consists of a flat surface of variable size with two electrodynamic exciters mounted on the back (Figure 1). Audio signals are fed to the exciters such that the surface acts as a loudspeaker membrane. The physical properties of the surface material lead to a certain radiation characteristics. The directional effect is further emphasised by crosstalk-cancelled stereophony, which creates a large auditory canvas that is much wider than the projection panel (see section III). The effect occurs at a central listening position, facing the device from a distance of a few ten centimetres up to about one and a half metres. As soon as the sweet area is left, the spatial auditory image collapses and sound seems to be restricted to the panel surface. Although both exciters are mounted to the same surface, their signal differences remain intact to a sufficient extent for providing the effect of crosstalk-cancelled projection.

A. Physical construction

Promenade has been prototyped with two materials and in two sizes so far. The first series uses finnboard, a lightweight type of cardboard, with the dimensions 70 by 100 centimetres (portrait) and 3 millimetres thickness. The second, more compact series is made of screenboard, a more dense type of cardboard with white cover layers, 43,5 by 64 centimetres and 2 millimetres thick. Other promising materials that have been explored include ply wood of a few millimetres thickness, and KAPA, a compound material combining a polyurethane foam core with cardboard cover layers.

The smaller prototype uses Visaton EX-60 R exciters (60 millimetres in diameter), while the larger is equipped with the more powerful Visaton EX-80 S (square shape, 80 millimetres edge length). White flat cable serves for feeding the exciters and for suspending the device. Both the exciters and the cables are glued to the panels. The projection devices are fed by small standard amplifier boards installed separately.

In the installation, the projected signals were generated from a BeagleBone Black computer board equipped with the Bela audio cape.[1] Sound programming and convolution for crosstalk cancellation was performed in SuperCollider.[2] As cancellation filters, the so-called BACCH filters by Edgar Choueiri [1] were taken from the jconvolver convolution engine bundle by Fons Adriaensen.[3]

B. Sound installation and artistic context

The sound projection device has been conceived as part of a sound installation within the arts-based research project *Promenade. Anekdote, Alltag, Abbild (Promenade. Anecdote, everyday life, image),* hence its name.[4]

Seven playback canvases (five in a later version, Figure 2) were distributed in the exhibition space, each suspended from the ceiling. The projection surface was left blank, without any further visual additions. Each of the devices played back a soundscape composition based on field recordings. Listeners were invited to explore and 'visit' the different soundscapes by moving in space, approaching the projection devices one by one and 'entering' the recorded space by positioning themselves in the sweet area where the crosstalk-cancelled projection effect unfolds.

The installation seeks to combine a global sonic environment that reflects the simultaneity of seemingly unrelated structures, processes, and realities around us, with the singularity of particular spots in space that, when located, allow for an exclusive access to individual layers of this environment. The crosstalk-cancelled projection technique restricts these spots to clear-cut boundaries but on the other hand it potentially 'elevates' each layer's spatiality as a sublime, temporary overlay to the global environment.

Figure 2. *Promenade* installation, SMC 2018, Limassol, Cyprus.

The installation *Promenade* serves as an experiential starting point for further investigations into the fields of anecdotal music, narrative identity construction and strollology. All of these areas may be reflected in conjunction with aesthetic experience in our everyday life auditory surrounding and its technical mediation. Departing from a workshop within the project *Promenade,* a publication is currently prepared that summarises the described approach and the ongoing process.

III. Transaural crosstalk-cancelled sound projection

A. Working principle

Transaural crosstalk-cancelled sound projection, often abbreviated XTC, has been developed for playing back binaural signals, which normally require headphones, on a pair of loudspeakers (see, e. g., [2, pp. 283–326]). While headphones isolate the sound propagation paths from each of the earphones to the corresponding ear, respectively, loudspeaker projection involves transaural crosstalk: the signal radiated from the left ear also reaches the right one and vice versa. However, the playback of binaural signals requires the separate presentation of each channel to the corresponding ear.

Transaural crosstalk cancellation is based on time-shifted anti-phase signals emitted from the respective opposite loudspeaker that shall arrive at the near-side ear at the same time as the unwanted crosstalk signal from the far side speaker in order to cancel it out. Obviously, the cancellation signal causes a crosstalk at the far side ear, too, which in turn is eliminated by another correction signal emitted from the opposite speaker. Several iterations of this process lead to the design of a recursive filter network that ultimately seeks to ensure an isolated projection of the ear signal pair [3].

[1] http://bela.io
[2] http://supercollider.github.io
[3] http://kokkinizita.linuxaudio.org/linuxaudio/
[4] https://esc.mur.at/de/projekt/promenade-anekdote-alltag-abbild

B. Binaural vs. stereophonic signals with XTC

As mentioned above, transaural crosstalk cancellation has been developed for the projection of binaural signals on loudspeakers, that is, ear signals involving interaural time (ITD) and level differences (ILD) as well as elevation-dependent spectral cues (see [4]). Nevertheless, XTC filters are also used for the playback of ordinary loudspeaker stereophonic signals. The latter application is known as *Ambiophonics* in consumer entertainment contexts [5].

Experimental findings of the *Promenade* project with respect to stereophonic signals are discussed below (Section V).

IV. Flat panel loudspeaker projection

The excitation of flat, solid matter surfaces as an alternative loudspeaker working principle has been investigated during the past decades (see, e.g., [6]). Major advantages are the appealing form factor requiring dramatically reduced mounting depth, and the potential to create diffuse sound fields due to the low directivity even at high frequencies.

One important factor for exciting flat panels is the mounting location of the exciter. Depending on the application, the physical properties and the size of the panel, a basic question is whether to exploit or to avoid the direct excitation of predominant modes. Often, an out-of-centre location avoiding low integer partition of the surface dimensions is recommended for a more balanced response. In *Promenade,* exciters were mounted according to the golden section as an experimentally found trade-off of acoustic and visual qualities.

Apart from the construction of high-end loudspeakers, electrodynamic excitation of manifold objects for sound projection is frequently applied in custom presentation media and in sound art (see, e.g., *SHHH SHHH SHHH*[5] by Amanda Dunsmore, 2008, or *Touched Echo*[6] by Markus Kison, 2007). As described above, the exploration for *Promenade* departs from the latter. In such realms the presented sound material is usually created or adapted especially for the specific playback medium. Consequently, the projection device does not need to meet the same criteria as general-purpose loudspeakers, especially with respect to linearity and neutrality, because such deviations can be corrected in the adaption process. This largely facilitates achieving a beneficial result with relatively low-cost equipment.

A special technical issue of the *Promenade* device is that only one panel is used for projecting two audio channels at the same time. The crucial question is whether the crosstalk cancellation cues are propagated to a sufficient extent for achieving a similar effect as with two small-spaced loudspeakers or whether the inter-channel crosstalk in the panel degrades the cues too much. A similar technique has been investigated for the application of Wave Field Synthesis (WFS) using even more than two exciters on the same panel, coined Multi-Actuator Panels [7]. Later, the method has been generalised for creating multiple sound sources with a single panel [8].

[5] http://www.lit.ie/Dunsmore/linz/kunstraumwindows.html
[6] http://www.markuskison.de/touched_echo.html

The experimental artistic context of *Promenade* did not pursue a scientific approach to such a technique for transaural sound projection. In multiple iterations, the used combination of materials, dimensions and mount points of the exciters have been adjusted such that the desired effect could be achieved sufficiently well. A validation by acoustic measurements, however, has not yet been carried out.

V. Discussion

The presented sound projection device shows an appropriate performance within its frame of origin, the *Promenade* arts-based research project. There, it allows for experiencing the often surprising, sometimes uncanny effect of crosstalk-cancelled sound projection and supports the exploration and reflection of aesthetic experience in multi-layered spatial auditory environments.

Beyond the attached context, the principles exploited in *Promenade* may be applied to other use cases. In the following, some findings and possible directions are discussed, such as the use of stereophonic signals, the integration with visual presentations, extensions towards interactive applications, and other form factors, materials, or objects.

A. Crosstalk cancellation of stereophonic signals

When developing the *Promenade* sound installation with field recording material and XTC projection, it was found that coincident stereophonic signals involving only level differences but none of phase provide a more consistent and stable spatial image than recordings in spaced omnidirectional stereophony that is based on phase differences. There is no well-researched explanation for this observation. It is likely that both the listening conditions and the projection device are not optimal for propagating phase differences to the listener's ears. In particular the fact that both exciters are working on the same projection surface may degrade the fidelity of phase reproduction, although a similar observation was made using two separate loudspeakers.

Overall, the discovery that intensity stereophony is well suited to span the extremes of the projection base – from ear to ear using headphones, nevertheless crossing the listener's head, from speaker to speaker in an ordinary stereophonic setup, and along a well externalised, virtual stereo base of about 150 degrees opening angle using XTC projection – is most valuable for the production of appropriate material. Following from that, field recordings were carried out in mid/side stereophony, which allows for further benefits with respect to adaptation and interaction (see subsection V-C).

B. Visual integration

In the *Promenade* sound installation, major emphasis was put on listening to mediated auditory environments. Therefore, the metaphor of a canvas was picked up that could have shown a landscape but instead projected a soundscape while remaining visually 'silent'. However, the surface of the projection medium may be easily used for a seamless integration of visual and auditory material. In the simplest case, the cardboard material may be printed or laminated. The backside may be used as well, restricted by the mounted exciters.

Alternatively, the panel may serve as a video projection screen, including three-dimensional technology involving polarised or shutter glasses. The projection angle and distance have to be envisaged such that the spectator is not disturbing the projection but may still position herself appropriately with respect to crosstalk-cancelled sound projection.

Using stereophonic signals, auditory events can be evoked for the listener anywhere in the horizontal dimension within the borders of the canvas and beyond. Therefore, horizontal audio-visual congruence is well achievable, for example, for annotating visual items in maps or alike. Psychoacoustic effects may be exploited for suggesting vertical correspondences. Depending on the stiffness of the projection panel, the location of the exciters on the panel is audible in certain frequency bands. In this case, a second pair of exciters may serve as a panning means in the vertical dimension, at least to a certain extent.

Generally, any visual presentation that favours a fixed, frontal and relatively close viewpoint can be expected to work well in combination with a *Promenade* like sound projection device.

C. Interactivity, tracking, sensors

From the beginning, the *Promenade* device has been conceived with an additional optical distance sensor such that the playback can react interactively, depending on whether, how close and for how long a listener is present in front of the canvas. For the installation versions involving several projection panels, actual use of the sensors has been dismissed so far. It turned out to be more desirable in the first place to explore the field of sweet spots distributed in space which, when being entered, change the listening experience dramatically. This effect may be regarded as an inherent ‘interactive’ behaviour of the projection dispositif, although no technical reactivity is involved.

Nevertheless, involving the sensor bears several interesting use cases. If multiple projection panels are present, the sensors may serve as a simple tracking device in order to activate or emphasis one or more panels rather than others, both in a decentralised (each panel controls itself without influencing others) or a centralised manner. Furthermore, the spatial arrangement of sound material could be extended towards an enterable, interactive radio play. Different individual listeners cannot be distinguished, though, which would require more sophisticated tracking means.

The sensor information may also affect the projection technique itself. In the *Promenade* installation, gradual transitions take place from monaural signals (centered in the projection panel) over ordinary stereophony (located roughly in between the exciter positions, thus a partition of the panel) to crosstalk-cancelled projection. As mid/side-stereophony is used, modulating the contribution of the side signal allows for further controlling the perceived width of the resulting sonic image. Such transitions may be triggered or emphasised by evaluating the listener’s distance or overall presence via the sensor.

D. Other form factors, materials, or objects

Theoretically, combining two-channel electrodynamic excitation of solid matter with transaural crosstalk cancellation is possible with various objects and materials of any form or extension. For example, architectural prototypes, engineering models or basic commodity might be enhanced by an integrated auditory layer that invites for spatial immersion.

The determining factors remain those that have been consulted while conceiving *Promenade:*

- the physical properties of the excited material with respect to sound propagation,
- appropriate mounting locations for the exciters and fitting form factors of the object and the exciters, and
- sufficiently isolated radiation of both channels for achieving the desired effect of crosstalk cancellation.

VI. Conclusion

In this paper, the *Promenade* spatial sound projection device has been presented, along with its originating context, a sound installation as part of the arts-based research project of the same name. The device has been presented as combining the two techniques of transaural crosstalk cancellation for inviting a rich spatial auditory experience with the excitation of solid objects for an unobtrusive appearance, differing from that of ordinary loudspeakers. The presentation form of a canvas, though not visually occupied in the original installation, invites for manifold applications that involve audio-visual correspondence. Further extensions include the use of distance sensors for interactive functionality, both with respect to the presented narrative and to the evoked spatial impression.

Acknowledgements

The research project *Promenade* is supported by the Vice Rector’s office of research at University of Music and Performing Arts Graz, the Cultural Office of the City of Graz, and ESC Media Arts Laboratory, Graz.

References

[1] E. Y. Choueri, “Optimal crosstalk cancellation for binaural audio with two loudspeakers,” 2010. [Online]. Available: https://www.princeton.edu/3D3A/Publications/BACCHPaperV4d.pdf

[2] B. Xie, *Head-Related Transfer Function and Virtual Auditory Display*. Plantation: J. Ross, 2013.

[3] R. Glasgal, “360° localization via 4.x RACE processing,” in *123rd AES Convention*. AES, 2007.

[4] J. Blauert, *Spatial Hearing*. Cambridge Mass./London: MIT Press, 1997.

[5] R. Glasgal, “Ambiophonics: The synthesis of concert-hall sound fields in the home,” in *99th AES Convention*. AES, 1995.

[6] G. Bank and N. Harris, “The distributed mode loudspeaker – theory and practice,” in *13th Microphone & Loudspeakers AES Conference*. AES, 1998.

[7] M. M. Boone, “Multi-actuator panels (MAPs) as loudspeaker arrays for wave field synthesis,” *Journal of the Audio Engineering Society*, vol. 52, no. 7/8, pp. 712–723, 2004.

[8] M. C. Heilemann, D. Anderson, and M. F. Bocko, “Sound-source localization on flat-panel loudspeakers,” *Journal of the Audio Engineering Society*, vol. 65, no. 3, pp. 168–177, 2017.