Audio Engineering Society

Convention Express Paper 4

Presented at the 153rd Convention
2022 October

Immersive Personal Sound Using a Surface Nearfield Source

Jukka Linjama[1] and Vesa Välimäki[1,2]

[1]*Flexound Systems Oy, Espoo, Finland*
[2]*Acoustics Lab, Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland*

Correspondence should be addressed to Jukka Linjama (`jukka.linjama@flexound.com`)

ABSTRACT

This paper discusses sound reproduction using a *surface nearfield source* (SNS), which is categorized between headphones and loudspeakers providing also a natural audio-tactile augmentation to the listening experience. The SNS can be embedded for example in the headrest as a personal sound system. In this sense it has similarities to headphones, but there is no need to wear a device. The SNS has also several advantages compared to loudspeakers, such as suppressed room effect and enhanced bass perception. Differences and similarities of the SNS approach with open and closed headphones, mobile device speakers, and regular loudspeakers are itemized. The SNS implementation is applicable to e.g. movie theater couches and car seats.

1 Introduction

Audio industry has developed various sound reproduction techniques for providing a satisfying listening experience. This challenge has been addressed mostly with two approaches: using headphones for individual experience, or using loudspeakers for shared or public audio reproduction in the acoustic farfield. People are increasingly using mobile devices, pushing the envelope of nearfield listening. Improved hands-free speakers and signal-processing techniques on these devices have prompted people to watch movies or conduct remote meetings without headphones.

Following this trend, wearable accessory speaker products have emerged, such as neck bands (e.g. [1]) or sound pillows with small speakers (or headphones) placed inside. Unfortunately, these do not fully solve the problem of disturbing others nearby, and their low-frequency response is very limited. A response to this lack of low-end is offered by those products with separate vibrating devices for personal bass sound reproduction, either installed in a seat, or as wearables. In practice, none of these solutions for personal nearfield listening are able to provide full frequency range audio without combining loudspeakers or headphones with separate vibration devices (e.g. [2]).

Human hearing is not a matter of ears only. Listening is inherently multisensory: it involves cues from other senses as well. Vibrotactile sensing is part of the listening experience [3]. Both our skin and our eardrums sense sound in the similar way – through mechanical vibrations – either via sound waves in the air,

Fig. 1: Surface nearfield source, or SNS, as a headrest yields a strong tactile bass and stereophonic personal soundscape. A HUMU Augmented Audio Cushion device by Flexound is used as an example here [6].

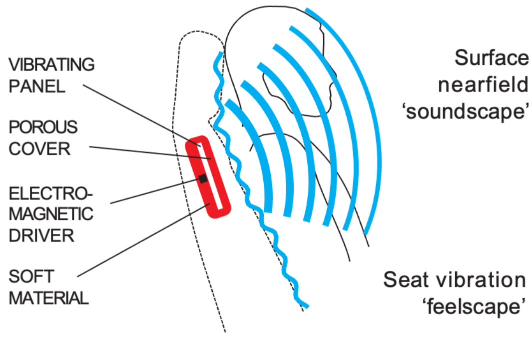

Fig. 2: Audio-tactile SNS speaker implementation in a seat with embedded Elastic vibrating element, EVE.

or through direct contact with a vibrating surface. It has been shown previously that vibration in the seat of the listeners enhances the music experience [4]. In addition to this, listening experience is also unconsciously and heavily affected by the visual environment ([5], p. xxi).

This paper discusses an alternative approach for sound reproduction without headphones, which goes beyond the limits of farfield and nearfield listening. This approach is based on a Surface Nearfield Source (SNS) that enables vibrotactile perception to be part of the experience. In terms of privacy, the SNS is categorized between the personal playback of headphones and the public playback of loudspeakers. Fig. 1 shows an example use case of an SNS implementation for personal sound.

The paper is organized as follows. Section 2 describes the implementation of an SNS and its properties as a sound reproduction system, Section 3 discusses different sound reproduction categories, Section 4 compares the SNS with headphones and loudspeakers, and Section 5 concludes.

2 Surface Nearfield Source

This section discusses the implementation of the SNS and its use as a personal sound reproduction system.

2.1 Farfield, Nearfield, and Surface Nearfield

Technically, *farfield* means a condition where the sound source is small compared to the wavelength and far away (several wavelengths) from the listening location. If these conditions (of a freely propagating spherical or plane wave) are not met, a more complex geometric *nearfield* behaviour is expected. In a closer proximity to a nearfield source, there is a region called hydrodynamic nearfield [7]. This volume in space extends a fraction of the wavelength away from the source surface. For relatively flat sources it may be called the *surface nearfield*, the term used in this paper for SNS.

2.2 Implementing a Surface Nearfield Source

Sound reproduction with SNS can be implemented using an embedded structure that we call an Elastic Vibrating Element (EVE). This is conceptually a "speaker element", but is more than that as EVE serves as a distributed audio-tactile source together with the structures surrounding it. EVE consists of a vibrating soundboard panel with an electromagnetic inertial driver inside soft, porous layers and a material cover, as shown in Fig. 2. The use of soft material distinguishes it from conventional flat panel speakers that are based on a directly radiating flexural wave field plate [8]. This EVE implementation of the SNS provides two major advantages:

- The EVE transmits both the vibrations and the full frequency range audio while being comfortably in contact with the listener;

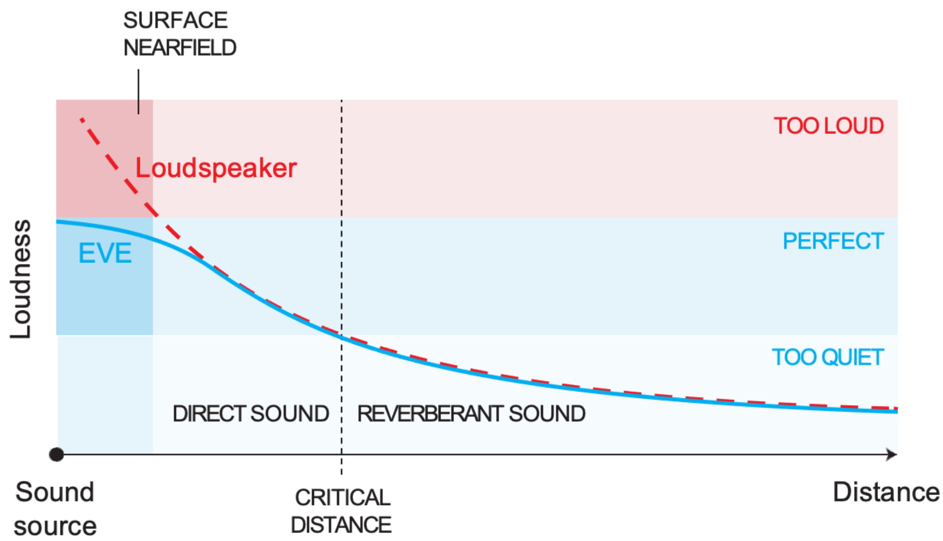

Fig. 3: Schematic illustration of Loudness vs listening distance in acoustic space: farfield, nearfield, and surface nearfield regions of a compact sound source. The critical distance defines the room nearfield where the direct sound dominates the reverberant sound in farfield. EVE is a distributed surface sound source, which enables close listening in its surface nearfield.

- The EVE can be integrated into a larger structure (such as a seat backrest) to transmit vibrations to a wider surface area, for a strong nearfield bass sound that does not radiate into the farfield, due to its dipole radiation characteristics.

An interesting paradox lies at the heart of the EVE's success in creating a balanced audio image, as part of the acoustic and mechanical signal from the vibrating panel is lost when it passes through the soft porous material layers. This mechanical damping is easily interpreted as causing "waste" of the acoustical power but paradoxically is the key enabler and foundation for successful SNS speaker solutions. When properly implemented, this kind of lossy "acoustic semiconductor" structure works in favor of the listening experience, because it evens out the spatial and frequency response characteristics of the surface nearfield, without compromising the treble signal too much.

2.3 Properties of SNS as Sound Source

An advantage of the SNS is that can be listened to at a very close distance without the sound being heavily localized, as shown in Fig. 3. The perceived sound level increases, when the listener is in close contact with the surface, but it does not get too loud, as with regular loudspeaker elements with a small distinct radiating surface. This is also illustrated in Fig. 3.

In practice, the SNS speaker is located close to the ears and located behind the head. Despite the source being behind the user, the listening experience is close to headphone listening, and in fact goes beyond that as it surprisingly resembles that of loudspeakers in a room. The way that the audio image wraps around creates a compelling binaural illusion of the surrounding acoustic space. This is partly because of the distributed source characteristis of EVE, and partly due to a fortunate side-effect of the inaccuracy of human directional hearing, when it comes to distinguishing between sound coming from the front or the back. This is known as the front-back confusion [3]; a phenomenon illustrated in Fig. 4 that occurs easily when the listener is not moving his/her head to support directional hearing.

Lateral (left/right) sounds are easier to pinpoint. The stereo SNS configuration in Fig. 4 in which two EVEs

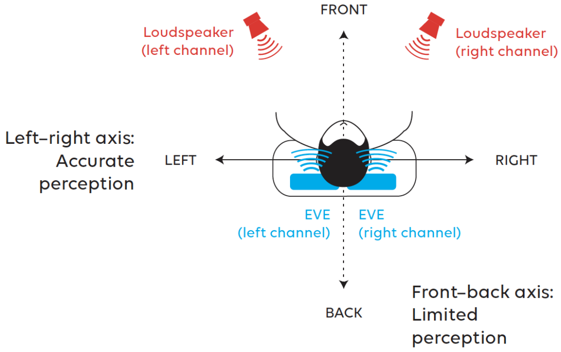

Fig. 4: Spatial perception characteristics of hearing in stereo sound reproduction. Lateral (left-right) direction perception is much more accurate than distinguishing sounds in front-back directions.

are used, one on each side of the head in the headrest, provides this information so that it is perceived effortlessly, despite the sound coming from the back. This is another paradox of the surface nearfield implementation: although the sound is played behind the user, in most listening contexts this does not present a problem for localization or clarity. The same applies to headphones, where sound image may be localized inside-the-head, without completely destroying the spatial listening experience.

3 Sound Reproduction Categories

This sections briefly discusses the different methods of sound reproduction, attempts to categorize them in terms of the listening context, and compares their main features. The goal is to highlight how the SNS shares attributes from both headphones and loudspeakers and how it differs from them.

3.1 Main Reproduction Technology Solutions

Sound reproduction takes place in a listening context, and regarding the degree of privacy (personal or shared experience), extreme solutions are headphones and loudspeakers. How does the SNS position itself between these? Table 1 attempts to clarify this by comparing four distinct sound reproduction categories, which are provided by headphones, regular loudspeaker, small loudspeakers, such as those of a mobile phone or a tablet, and the SNS. In terms of privacy, the headphones are the best solution for personal listening. Loudspeakers, on the other hand, are mainly targeted to public listening. Mobile devices and the SNS are both midway between personal and public, as they are primarily suitable for a single user but the sound leaks to the environment.

The main difference between regular and mobile loudspeakers is in the placement, since the former are listened to in the farfield whereas the latter ones are used in the nearfield, where minor movements lead to large changes in the response. In headphones, the sound is fed directly to the user's ears, and they are thus not located in either far- or nearfield. The SNS works in the surface nearfield and connects best to the listener's ears, when the user's head is close to the surface, and in practice, SNS is located behind the head.

An additional advantage of the SNS is that vibration is transmitted to the listener's body, thus providing vibrotactile perception [6], or feel, as implied in Table 1. With regular loudspeakers, tactile sensation is only possible in the presence of very loud low-frequency sound [9]. With moderate listening volumes, conventional headphones or small loudspeakers used in mobile devices do not provide feel either.

4 Comparison of Reproduction Technologies

Next, the features of headphones and loudspeakers are compared with the SNS, in order to further clarify the possible role of SNS approach in sound reproduction. This time we consider the two main headphones types, open and closed, where "open" refers to open-back headphones with little acoustic isolation, such as most intra-concha (button-type) and some circumaural (over-ear) headphones. All in-ear earphones, which are inserted in the ear canal as earplugs, and other isolating headphones, which considerable suppress external sounds, are "closed". Loudspeakers are again classified in two categories in Table 2: small nearfield and regular farfield loudspeakers. Table 2 summarizes the following comparisons between these listening use cases.

4.1 Personal vs. Public Listening

Personal playback is perfectly implemented with closed headphones, which isolate the listener from ambient

Table 1: Sound reproduction solutions and some of their basic attributes.

Technology solution	**Headphones**	**Surface nearfield source (SNS)**	**Mobile device loudspeakers**	**Loudspeakers**
Privacy	Personal	Personal/Public	Personal/Public	Public
Placement	On ears	Surface nearfield	Nearfield	Farfield
Audio-tactile feel	Audio only	Large surface area vibration	Audio only	Low frequencies through air at high SPL

Table 2: Descriptive comparison of sound reproduction technologies (⋆⋆⋆ = excellent/much, ⋆⋆ = good/some, ⋆ = a little, - = no/non-existent).

Feature	Closed headphones	Open headphones	Surface nearfield source	Small nearfield loudspeakers	Farfield loudspeakers
1) Personal playback	⋆⋆⋆	⋆⋆	⋆⋆	⋆	-
2) Public playback	-	-	⋆	⋆	⋆⋆⋆
3) No need to wear	-	-	⋆⋆⋆	⋆⋆	⋆⋆⋆
4) Max loudness	⋆⋆⋆	⋆⋆⋆	⋆⋆	⋆⋆	⋆⋆⋆
5) Full frequency range	⋆⋆⋆	⋆⋆⋆	⋆⋆⋆	⋆	⋆⋆⋆
6) Spatial immersion	⋆	⋆	⋆⋆	⋆	⋆⋆⋆
7) Directional 3-D sound	⋆	⋆	⋆	-	⋆⋆⋆
8) No room reflections	⋆⋆⋆	⋆⋆⋆	⋆⋆⋆	⋆⋆	-
9) Audio-tactile augmentation	-	-	⋆⋆⋆	-	-
10) Tactile bass extension	-	-	⋆⋆⋆	-	⋆

noise and which do not leak the sound to the surroundings, see Table 2, line 1. Open headphones are also good for personal use, although they leave some room for improvement in a noisy environment, as their ambient noise attenuation is limited. The SNS can be used for personal sound in very much the same way as open headphones, as indicated on line 1 in Table 2. Minor downsides of the SNS are that some sound will leak to outsiders and there is no acoustic isolation. Small mobile speakers can also be used for personal sound, but they additionally lack the good bass response.

Farfield loudspeakers are the best option for public playback of sound but non-ideal for personal playback, as everyone in the same space usually hears everything. There are exceptions, such as directive monitor loudspeakers, which are targeted for personal use, but are not discussed here further. Small mobile device speakers are not powerful enough for public playback. Similarly, the SNSs do not radiate efficiently to the environment in the full audio range, as shown on line 2 in Table 2. In practice, mainly the high-frequency sound from the SNS can be heard from a distance. The SNSs still work better as regular speakers than any headphones.

SNSs are easy to use since there is no need to wear or hold anything, and are in this respect regular loudspeaker, see line 3 in Table 2. On the other hand, headphone users frequently complain about the discomfort caused by the fit, size, or cables [10].

4.2 Loudness and Frequency Response

In terms of maximum loudness, the SNS is comparable to small mobile speakers: very high SPL at low frequencies is hard to achieve without creating excessive vibration. Modern headphones and larger loudspeakers are better at this, as indicated on line 4 in Table 2.

Nonetheless, the SNS can provide high-quality sound on the full audio range in a way similar to headphones and proper hi-fi loudspeakers. Thus, the SNS is a better sound reproduction solution than the tiny speakers used in smart phones, tablets, or laptops, which are lacking in bass, see line 5 in Table 2.

4.3 Spatial Immersion and Directional Sound

Compared to headphones, an SNS stereo implementation provides an immersive soundscape that is not localized inside the head, as indicated on line 6 in Table 2. With a good stereo recording, a very natural illusion of the surrounding acoustic space can be achieved.

Accurate directional sound (2-D or 3-D) rendering in the SNS is fairly limited in the same way as in conventional headphones, see line 7 in Table 2. Headphones equipped with head-tracking and binaural signal processing would be much better, but are not discussed here. A preliminary study of implementing binaural audio processing with (farfield) head-related transfer function and cross-talk filtering to SNS indicates that the task is not easy, and requires head-tracking [6].

4.4 Room Effect and Nearfield Listening

Regular loudspeakers placed in the farfield create an audio image, which consists of the direct sound of the speakers as well as the reflections from the surfaces in the listening room. At a range further than the critical distance, the reflected sound in the room overrides the direct sound, with the sound level being mainly defined by the room reverberation time [11]. Enormous efforts have been put to deal with this challenge by the design of speaker element directivity, acoustic design of listening rooms [5], and by studying signal-processing techniques for loudspeaker-room equalization [12].

If the listener gets close enough to the loudspeakers, they are in the nearfield, where room reflections are weaker than the direct sound. In principle, nearfield listening provides a purer reproduction of the audio material, as the room effect is suppressed. In essence, the speaker placement and listener location are critical to getting the desired sound reproduction.

The advantage of nearfield listening—either with studio monitor speakers or headphones—is in the accuracy and clarity of the direct sound, which is unaffected by the room. However, if the sound source is compact, as with the treble elements in multi-way loudspeakers, the sound image in the nearfield is highly dependent on the distance to each ear: it may be localized at the source, and becomes overly loud if the ear is moved close to the source. With the SNS, the listener can enjoy the same reflection-free reproduction as with headphones, see line 8 in Table 2.

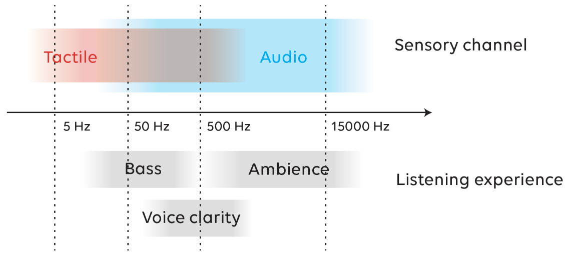

Fig. 5: Audio-tactile perception of sound. Useful frequency ranges of tactile and auditory sensory channels, and their relation to bass, ambience, and voice clarity enhancements in audio-tactile SNS implementation.

4.5 Tactile, Audio-tactile, and Auditory Perception

The SNS shines especially in its audio-tactile augmentation of sound, as indicated on lines 9 and 10 in Table 2. The role of tactile and auditory sensing of sound across the auditory frequency range is illustrated schematically in Fig. 5. In addition to the low bass that is in fact perceived primarily through vibration at frequencies below 50 Hz, the tactile channel also supports hearing at frequencies up to 500 Hz. This overlapping audio-tactile frequency range can help in adding (multisensory) clarity to the voice dialog, both for people with normal hearing and especially for the hard-of-hearing people.

Conventional headphones or loudspeakers cannot provide similar vibrotactile sensations, without using either additional vibration accessories, or ultimate loudness. The SNS implementation enables a rich but silent listening experience due to its multisensory nature. One constraint for tactile perception is that it has a more limited dynamic range than hearing. This constraint can be taken into account in audio signal processing, by using dedicated “feelness control” dynamics compression on the tactile frequency band.

In addition to the audio-tactile benefits, the SNS approach contributes to pure auditory soundscape features in the frequency range above 500 Hz, as discussed above, and titled as ambience for brevity in Fig. 5.

5 Conclusion

This paper introduces and discusses the SNS, a new kind of surface nearfield speaker solution that brings more clarity and immersion to personal audio, by also experiencing sound through body contact. The SNS has similarities to headphones, as the direct sound energy dominates and thus the effect of room reverberation on the sound is greatly reduced. However, in contrast to headphones, the SNS does not require wearing anything, but can be enjoyed by being close to (or in contact with) the device. Paradoxically, the sound emanating from the SNS is usually localized around the listener even though the vibrating elements are located behind the user.

The core of the SNS-based audio implementation is the Elastic Vibrating Element, EVE, which provides full frequency range audio, together with vibration from a single source. In combination with appropriate audio signal processing, the EVE can be integrated into the seats or cushions to deliver an embedded, invisible loudspeaker system. It may be used either as a personal stand-alone listening solution, or with dedicated multichannel mixing, to augment conventional cinema or car loudspeaker systems with deeper immersion to the sound content.

References

[1] Imaizumi, K. and Haneda, Y., "Wearable sound reproduction system using two end-fire arrays," in *Proceedings of the Audio Engineering Society 142nd Convention*, May 2017.

[2] Polygon, "This chest-thumping wearable adds a surprising amount of emotion to gaming," 2017, https://www.polygon.com/2017/12/22/16811074/woojer-strap-vibrating-wearable-game-hardware (accessed: July 7, 2022).

[3] Pulkki, V. and Karjalainen, M., *Communication Acoustics: An Introduction to Speech, Audio and Psychoacoustics*, John Wiley & Sons, 2015.

[4] Merchel, S. and Altinsoy, M. E., "Auditory-tactile experience of music," in S. Papetti and C. Saitis, editors, *Musical Haptics*, pp. 123–148, Springer, Cham, 2018.

[5] Toole, F. E., *Sound Reproduction: The Acoustics and Psychoacoustics of Loudspeakers and Rooms*, Routledge, 3rd edition, 2018.

[6] Puomio, O., González Diaz, R., and Lokki, T., "Augmented audio with behind-the-head listening device," in *Proceedings of the AES International Conference on Immersive and Interactive Audio*, Mar. 2019.

[7] Fahy, F. J., *Foundations of Engineering Acoustics*, Elsevier, 2001.

[8] Heilemann, M. C., Anderson, D. A., Roessner, S., and Bocko, M. F., "The evolution and design of flat-panel loudspeakers for audio reproduction," *J. Audio Eng. Soc.*, 69(1/2), pp. 27–39, Jan./Feb. 2021.

[9] Takahashi, Y., Yonekawa, Y., Kanada, K., and Maeda, S., "A pilot study on the human body vibration induced by low frequency noise," *Industrial Health*, 37(1), pp. 28–35, Jun. 1999.

[10] Rane, M., Coleman, P., Mason, R., and Bech, S., "Survey of user perspectives on headphone technology," in *Proceedings of the Audio Engineering Society 152nd Convention*, May 2022.

[11] Howard, D. and Angus, J., *Acoustics and Psychoacoustics*, Routledge, 5th edition, 2017.

[12] Cecchi, S., Carini, A., and Spors, S., "Room response equalization—A review," *Appl. Sci.*, 8(1), paper 16, 2018, `https://doi.org/10.3390/app8010016`.