Journal on Multimodal User Interfaces (2022) 16:125–142
https://doi.org/10.1007/s12193-021-00378-8

ORIGINAL PAPER

Combining audio and visual displays to highlight temporal and spatial seismic patterns

Arthur Paté[1] · Gaspard Farge[2] · Benjamin K. Holtzman[3] · Anna C. Barth[3] · Piero Poli[4] · Lapo Boschi[5,7,8] · Leif Karlstrom[6]

Received: 30 March 2021 / Accepted: 12 July 2021 / Published online: 27 July 2021

Abstract
Data visualization, and to a lesser extent data sonification, are classic tools to the scientific community. However, these two approaches are very rarely combined, although they are highly complementary: our visual system is good at recognizing spatial patterns, whereas our auditory system is better tuned for temporal patterns. In this article, data representation methods are proposed that combine visualization, sonification, and spatial audio techniques, in order to optimize the user's perception of spatial and temporal patterns in a single display, to increase the feeling of immersion, and to take advantage of multimodal integration mechanisms. Three seismic data sets are used to illustrate the methods, covering different physical phenomena, time scales, spatial distributions, and spatio-temporal dynamics. The methods are adapted to the specificities of each data set, and to the amount of information that the designer wants to display. This leads to further developments, namely the use of audification with two time scales, the switch from pure audification to time-modulated noise, and the switch from pure audification to sonic icons. First user feedback from live demonstrations indicates that the methods presented in this article seem to enhance the perception of spatio-temporal patterns, which is a key parameter to the understanding of seismically active systems, and a step towards apprehending the processes that drive this activity.

Keywords Auditory display · Visual display · Spatial audio · Data representation

Contents

✉ Arthur Paté
arthur.pate@isen.fr

1 Univ. Lille, CNRS, Centrale Lille, Univ. Polytechnique Hauts-de-France, Junia, UMR 8520 – IEMN, Lille, France

2 Université de Paris, Institut de physique du globe de Paris, CNRS, 75005 Paris, France

3 Lamont Doherty Earth Observatory, Columbia University, Palisades, NY, USA

4 ISTerre, Université Grenoble Alpes, UMR 5275 CNRS, Université Savoie Mont-Blanc, IRD, IFSTTAR, Grenoble, France

5 Dipartimento di Geoscienze, Universita degli Studi di Padova, Italy

6 Department of Earth Sciences, University of Oregon, Eugene, OR 97403, USA

7 Institut des Sciences de la Terre Paris, Sorbonne Université, CNRS-INSU, ISTeP UMR 7193, 75005 Paris, France

8 Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy

1 Background

This article falls within the area of auditory display, or sonification, *the use of nonspeech audio to convey information* [1]. Contrasting with the dominant approach in science where information is traditionally visualized, auditory display mobilizes our auditory system to access and analyze data. The permanence and materiality of visualizations such as printed text, images, and graphics has helped settling vision as our main medium for accessing information and debating it [2]. However, auditory display has a tremendous potential thanks to the abilities of our auditory system, e.g. its ability to extract meaningful signals from very noisy environments or separate between multiple audio streams, or the possibility to hear and understand sonic cues without actively and intentionally listening. Auditory display has been described elsewhere (see [3,4] and the yearly International Conference on Auditory Display, or ICAD), and Sect. 2 will discuss further the differences and complementarities between sonification and visualization.

In particular, this article deals with the sonification (and visualization) of seismic data, reduced here to the temporal evolution of the velocity of a point at the surface of the Earth, as measured by a seismometer and to discrete events in these observations. The recorded seismic waves may originate from natural earth processes such as earthquakes, volcanic eruptions, debris flows, or from human-induced sources, such as the recorded vibrations due to traffic or construction works ; all superimposing upon the continuous Earth "hum" arising from e.g., ocean infragravity waves, and commonly refered to as ambient noise. These seismic waves "look like" sound waves (i.e. oscillations around a mean value), so that it is enough to speed them up[1] to make them fall into the hearing range (seismic waves typically range from a few mHz to a few tens of Hz). By the time when seismic waves were recorded on magnetic tape, audification was common practice to seismologists: playing the tapes as audio tapes at higher rotation speed enabled the seismologists to quickly scan the recordings and identify the location of seismic events [7]. During the Cold War, Speeth [8] and later Frantti and Levereault [9] studied how listening to audified seismograms could help discriminate between earthquakes and (legal or illegal) nuclear tests [10]. Quite forgotten during almost 30 years, audification of seismic data has lately regained popularity [11–18].

[1] This process, known as audification, is maybe the first sonification technique that was used: in 1878 the early technology of telephone was used to listen to nervous impulses in muscles [5], cited by Dombois [6]; in 1924 bat scream recordings were slowed down and listened to [7]. On the contrary, electromagnetic waves have too high frequencies to be audible, they need to be slowed down for tuning the frequencies down until they reach our hearing range.

2 Motivation of the work

In this section we present a short discussion on data representation. There are, among others, two dimensions of a data set that are commonly seen as relevant to represent, and indeed very useful for seismic and geophysical data: time evolution and spatial distribution. The tradeoff between time and space is nicely put by Gaver [19]: *A simple way to contrast listening and looking is to say that although sound exists in time and over space, vision exists in space and over time. Sounds are well suited for conveying information about changing events.*

It is now generally acknowledged [4] that sonification techniques are efficient in giving the listener a sense of the time evolution of the data. Audification in particular preserves and highlights the time patterns, although it scales those patterns up or down to reach a speed our auditory system can handle. Sonification is deemed as more efficient than visualization for temporal matters, as our human auditory system is more sensitive to subtle and fine changes in time than our visual system [20]. Traditional display methods in geophysics and seismology are however based on either static pictures showing a 2d-map where time information might be indirectly displayed (e.g., all events plotted on the same map with similar markers, and a color code reflects the event time), or on waveform display as data point series indexed with time and visually rendered as a 2-axis plot.

On the other hand, the spatial distribution of the data is often provided to the users through visual displays: the idea of mapping information about space on an abstract and reduced space as a 2d- or 3d-visual representation (map) has pervaded our daily practices to a very deep point. It is clear that our visual system shows better abilities to discriminate subtle changes in spatial patterns than our auditory system: Walker et al. [21] rely on Howard & Templeton [22] to state that our auditory system has *angular resolutions approximately 10 times more coarse than the eye across the sensorially richest regions* ; Deutsch [23] shows that localization is not used as a primary cue for grouping sonic events, but rather used *only when other supporting cues are present*, and Bregman [24, p. 75–79] reports several cases where the spatial cue is outperformed by the frequency cue in establishing auditory streams.

Yet we can hear and precisely track a source move around us [25], or finely discriminate between sound sources located at different positions [26,27]. Contrarily to our visual system, our auditory system detects sources in a 360°-area (vs. approximately 120° for vision) [20]. However, and to our knowledge, spatialization is not the most common technique that is used in the field of auditory display. For example, the index of the well-established reference known as the *Sonification Handbook* [4] has very few entries for the key-

words "3D sound", "spatialization", "Ambisonics", "WFS[2]", "VBAP[3]", "DBAP[4]", "binaural", or "HRTF[5]" (the latter 6 being emblematic spatial audio techniques) ; a quick scan through the articles and extended abstracts proceedings of ICAD 2018 (resp. 2019) shows that only 2 (resp. 1) papers use Ambisonic technique, 3 (resp. 3) articles use more than 2 loudspeakers, and 1 (resp. 3) uses HRTFs. However, sound spatialization is of course known to the community, and has been successfully used, e.g. for providing the blind with a sonic image of their environment [28], in the design of auditory alarms, e.g. for conveying the sense or urgency [29] , or for the auditory representation of geographical maps [30,31]. Walker et al. [21] sonified the data from a personal digital assistant, associating the sound's location to the hour in the day, and Brungart and Simpson [32] produced a sonification for pilots where interaural differences were linked to the pitch of an aircraft. Closer to the data used in this article, and with a clear artistic purpose, the valuable contribution of Dombois should be mentioned here, e.g. his work "Circum Pacific 5.1"[6] rendering audified seismic waveforms from 5 different seismic stations with 5 loudspeakers (configurations of stations and loudspeakers being both similar to the classical 5.1 speaker configuration) [33]. Using spatial audio is known to increase our capacities, for example stereo panning [34] and HRTFs [35] increased stream segregation against mono signals, but at the expense of cognitive load: In a matching task between visual and audio graphs, Bonebright et al. [36] observed that stereo stimuli are more time-consuming and difficult to handle than mono ones. Also, the design guidelines by McGookin and Brewster [37] encourage the use of spatialization to help distinguish between earcons.

The fields of auditory display or visual display in isolation may well provide convincing methods that precisely render both temporal and spatial aspects of the data. It seems however that the combination of both auditory and visual displays represents a direct way to efficiently expose the users to the spatio-temporal features of the data. Such combination attempts are already known in the literature, e.g. for the *Seismodome* shows [15,38,39], or in similar endeavors [14,40,41].

The contribution of this article[7] is to bring together auditory and visual displays, each one bringing spatial and temporal information. This can be seen as a step toward a unified display that gives meaningful insights into both space and time dimensions, in the present case as applied to seismic data. It is believed that combining all these aspects can enhance the quality of the data representation in several ways. First, issues of single-modality displays can be easily solved, e.g. when static plots with colors according to time quickly become unreadable when points superimpose, or when the angular/spatial resolution of the human auditory system is not high enough to detect subtle spatial patterns. Second, multimodal integration is known to enhance our perception abilities, in particular when the information conveyed by the individual sensory modalities are close in time and space, and semantically congruent [42]. Third, using displays or environments that stimulate several sensory modalities enhance the feeling of immersion and involvement in the scene that is represented [43,44], presumably leading to increased attention and acuteness. Fourth, as visual displays are accessible to the deaf and auditory displays are accessible to the blind, audio-visual displays should definitely be accessible to a larger population. Obviously, one of the conditions for meeting this enlarged accessibility is that each display (visual or audio) presents the core aspects of the data independently of the other display, so that a user exposed to the auditory display only doesn't miss important features of the data that are accessible only through vision, and conversely.

This article presents the audio-visual representation of three seismic data sets, covering different physical phenomena, time scales, spatial distributions, and spatio-temporal dynamics. Section 3 presents the data. Then Sects. 4 and 5 respectively present the implemented auditory and visual display techniques, for each data set. Finally, Sect. 6 discusses the choices that were made, partly based on user feedback, and gives future directions.

3 Data

3.1 Kilauea volcano

The eruption of the Kilauea volcano (Hawaii, USA) in the summer of 2018 marked the end of a summit eruption phase that began in 2008. Seismic data in the vicinity of Kilauea volcano represents one of the principal monitoring tools of the U.S. Geological Survey (USGS) Hawaii Volcano Observatory. Seismicity recorded by the USGS seismic network has been widely used to document changes in volcanic activ-

[2] Wave Field Synthesis.

[3] Vector-Based Amplitude Panning.

[4] Distance-Based Amplitude Panning.

[5] Head-Related Transfer Functions.

[6] Description available at http://floriandombois.net/works/circum-pacific.html. See also Dombois's "Surf" (2006) at http://floriandombois.net/works/acceleration.html and "Acceleration 2,200" (2003) at http://floriandombois.net/works/acceleration.html ; all URLs last retrieved June 16, 2021.

[7] Part of this work was presented as a demo (no paper issued) at the Computer Music Multidisciplinary Research (CMMR) Symposium in Marseille, France, in October 2019.

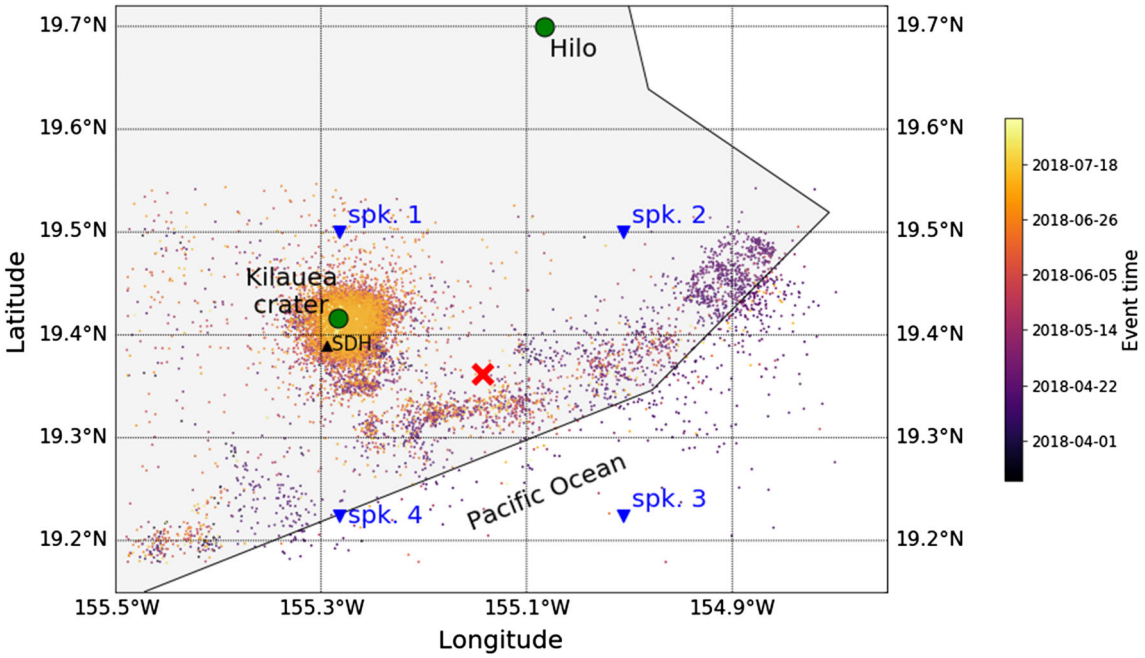

Fig. 1 Map view of the "Kilauea" data set. Selected events (magnitude < 1.5) are represented with dots whose color varies with event time, from blue (April 2018) to yellow (August 2018). The black upward pointing triangle represents the selected seismic station. Blue downward pointing triangles indicate the virtual position of the loudspeakers when playing the sounds (red cross indicates the center of the loudspeaker array). Green circles give landmarks to help the reader locate the area of interest: the crater of Kilauea volcano and the city of Hilo, on Hawai'i island

ity [45] as well as to inform models of the physical processes involved in subsurface magma motions [46]. We focus on a catalog of seismic events that record the spectacular summit caldera collapse sequence between late April and early August, 2018. Beginning on April 30th, seismicity migrated about 40 km down the southeast flank of Kilauea volcano towards where magma would eventually erupt and the frequency of seismic events rapidly increased in the vicinity of the summit caldera. After a series of explosive eruptions at the summit and initiation of voluminous magma effusion in the "Lower East Rift Zone", the eruption entered a phase in which a magnitude-5 earthquake occurred every day for over a month, accompanied by collapse events in the caldera that devastated the summit and destroyed the Hawaiian Volcano Observatory visitor center. Each major seismic event was preceded by a rapidly increasing number of foreshocks, followed by a brief "silence" (approximately 2 hours), and then a repeat. These earthquakes had no aftershocks, only foreshocks, which is uncommon for normal non-volcanic earthquakes. During the main collapse sequence, magma flow rate approximately 40 km away from the summit towards the primary eruption sites increased for several hours following the M5 summit collapse events, although magma flow itself proceeded aseismically [47].

A list of seismic events was obtained through the USGS webservices[8] as a catalog of 41,328 events, i.e. all events greater than magnitude 1.5 having occurred between April and August 2018 (5 months). For each of the events a 80-second recording was extracted from the trace of station called SDH (from the Hawaiian Volcano Observatory Network[9], sampling frequency 20 Hz), only the vertical component (motion perpendicular to the plane of the Earth surface) was kept, and the instrument response was removed. Figure 1 shows a map of the selected events, with a color code according to event time.

As stated above, the audiovisual, spatialized display of this data set aims at highlighting (a) the migration of seismic activity from the offshore to the crater, (b) the periodicity of the pattern of foreshock/main shock sequences, of approximate duration one day. It is expected that this display helps perceive the spatio-temporal patterns of seismicity, and possibly track down the migration of volcanic fluids, e.g. magma opening and circulating in fractures ("dykes" or "sills") [48,49], and landslides, slumps, and regular old frictional tectonically-driven earthquakes.

[8] Available at https://volcanoes.usgs.gov/volcanoes/kilauea/monitoring_kilauea.html.

[9] Information available at https://volcanoes.usgs.gov/observatories/hvo/.

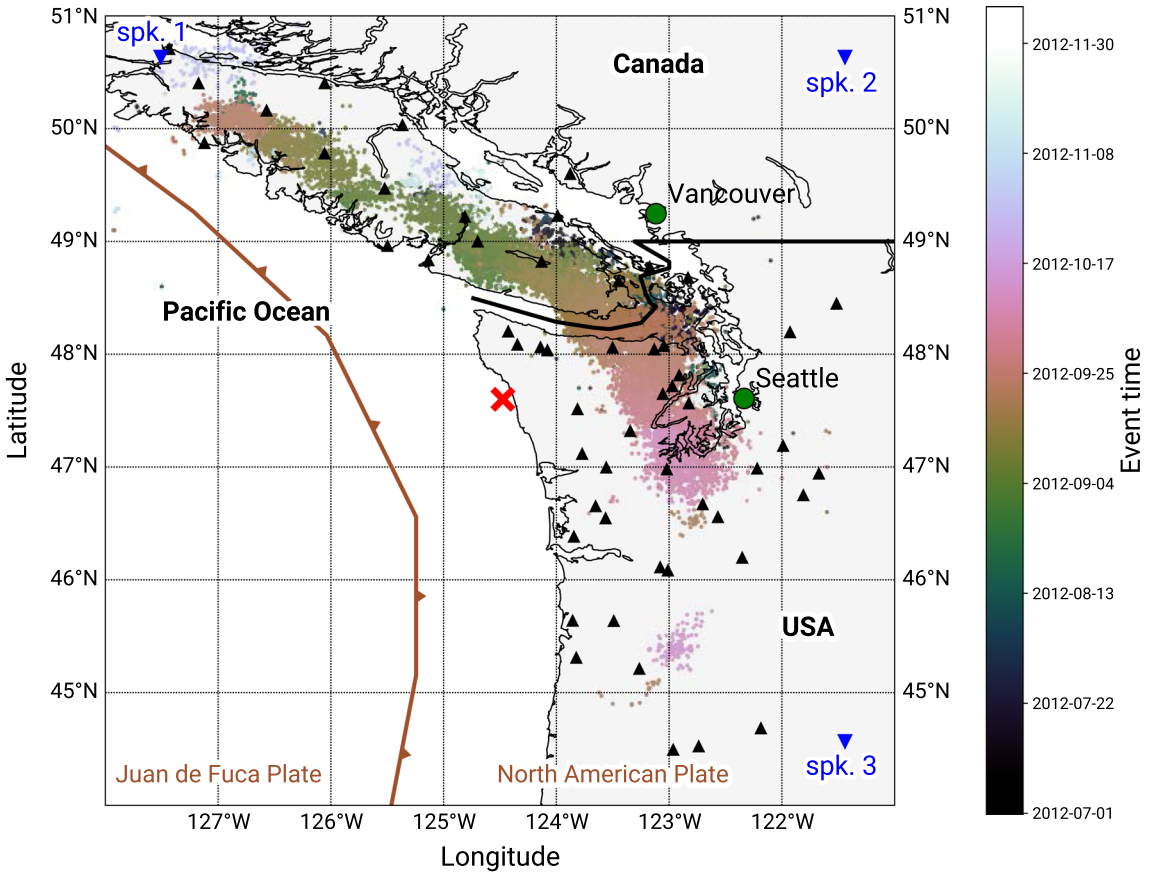

Fig. 2 Map view of the ETS data set. Tremor source locations in time are represented with dots whose color varies with event time, from black (July 2012) to white (December 2012). Each event corresponds to a 5-min segment of tremor emission, radiating seismic energy equivalent to a magnitude 1.5–2 earthquake. The black upward pointing triangles represent the selected seismic stations, with their names attached. Blue downward pointing triangles indicate the virtual position of the loudspeakers when playing the sounds (red cross indicates the center of the loudspeaker array). Green circles give landmarks to help the reader locate the area of interest: the cities of Vancouver (British Columbia, Canada) and Seattle (Washington, USA). Brown color is also used to indicate the trench line (solid line with triangle markers attached) and tectonic plate names

3.2 Episodic tremor and slip in the Cascadia subduction zone

The second data set is an example of tremor activity during an event of "Episodic Tremor and Slip" (ETS). The ETS event occurred in the Cascadia subduction zone (Pacific Northwest, USA and Canada) in October 2012. During such events, the Juan de Fuca tectonic plate (off the coast) dives below the North American plate, accommodating their continuous convergence. As during an earthquake, the plates slide, or "slip" against each other. In this case however, the plates move relative to each other much slower than during earthquakes. Such events of slow-slip are often called "slow earthquakes". Episodes of slow-slip of this kind are accompanied by very weak, low-frequency seismic emissions called "tremor" [50]. The location of the tremor's source in time allows to track the slow-slip front propagation, that is the progressive migration of the unzipping of the plates' interface, and of the fluid pressure pulse that is thought to be driving it in a complex feedback loop (see Frank et al. [51] for the same phenomena in the Mexican subduction zone).

Contrarily to earthquakes' seismic signals, tremor is a weak, non-impulsive, minutes- to hours-long signal, seemingly emitted from a broad epicentral volume, and thus less evident to separate from noise as a discrete event. Tremor source location can however be achieved by cross-correlating segments of the tremor signal's envelope recorded at different locations, and detecting the time lag—and thus the source-receiver distance—associated with each station [52]. Each "event" in our dataset thus corresponds to the most probable source of a segment of the tremor signal. This method of detection provides us with discrete source locations in time, but in reality the fine scale spatial structure is lost in order to more easily capture the large time- and space-scale variations of activity. Therefore, this dataset will be best used to highlight tremor patterns and apprehend the processes that drive

them on day-long, kilometric scales. Tremor source locations are mapped as dots in Fig. 2, with a color code according to event time.

Five months of tremor activity were investigated. A catalog of 33,880 events (tremor source locations) recorded by 55 stations from July 2012 to January 2013 (Pacific Northwest Seismic Network[10], sampling frequency 20 Hz) was compiled from the automatic tremor detection set up by the Pacific Northwest Seismic Network[11]. Each event was associated with an exploitable seismogram, recorded on a nearby station. Many seismograms show high-frequency (maybe anthropogenic) noise, or click-like artifacts, therefore we chose the first exploitable record in an array of stations closest to the event. Each six-minute-long seismic recording was limited to its horizontal East-West component, band-pass filtered between 1 and 8 Hz, as it is the frequency band at which the tremor signals are observed.

Before, during and after an ETS event, tremor activity exhibits complex but structured spatial and temporal patterns, that are uneasy to apprehend with traditional visualization tools. The audiovisual, spatialized display of tremor activity aims at highlighting (a) the stark change of activity style as the ETS event starts, and (b) the complex migration of activity during the event. By better perceiving these patterns of activity, it is easier to form a mental picture of the evolution and extent of the active slip and fluid processes that drive the activity.

3.3 Pollino earthquake swarm

The third data set is a continuous record of ambient seismic noise in Pollino National Park, South of Italy (North-East of Calabrian subduction zone). This micro-seismic activity, made of thousands of small earthquakes, is called an "earthquake swarm". The swarm sequence started in October 2010 and continues since then (maximum magnitude reached within this sequence is 5.0 on October 25th, 2012).

The recording of micro-seismic activity on April 29th, 2015 was downloaded [53]. The signals (24 hours, sampling frequency 20 Hz, vertical component) are continuous recordings made by 7 stations from the Y4 network[12]. This day was chosen because of its relatively low seismic activity. As micro-seismicity is observed, we don't rely on any event catalog, as it would be non-relevant (detected events are outside of the considered geographical area, or outside of the observed time window). That is why we work with a so-called "continuous recording" i.e., not segmented according to any available meta-data. Figure 3 shows a map of the selected stations.

As in the first two data sets, it is hypothesized by seismologists that the spatio-temporal patterns within the swarm (i.e., migration of micro-earthquakes) is meaningful to observe if targeting the understanding of geophysical mechanisms. The spatialized auditory display aims at making these patterns easier to grasp for the user. Besides, as seen in Fig. 4 representing the recording of one seismic station during the 24 hours of April 29th, 2015, swarm recordings are visually very similar to noise, making visual analysis or the use of analysis algorithms based on vision ineffective. The human auditory system, however, is powerful and efficient when relevant signals have to be separated from noise (see e.g. Bregman [24] for issues in segregating audio streams). The secondary aim of this auditory display is to test whether human listeners are able to tell relevant seismic information from noise, including anthropogenic noise (i.e., human-induced vibrations through road traffic for instance).

4 Auditory display

4.1 Audio setup

The methods described in this article were implemented, and audiovisual scenes were rendered in two different places (Lille, France; Marseille, France), and this paragraph gives an overview of the setup, providing guidelines for the reader. In both locations the room was a square with a surface of roughly 70 m^2. Four loudspeakers were placed as vertices of a square of side length approximately 4 m, in the middle of the room. The self-powered, amplified loudspeakers were either *Alpha 40* by *Focal* (Lille) or *8020* by *Genelec* (Marseille), connected to an audio interface that was either a *Scarlett 18i20* by *Focusrite* (Lille), or a *Rubix 24* by *Roland* (Marseille). Seismic signals were sonified using a *Python* script, resulting in four tracks (1 for each loudspeaker), and the digital audio workstation *Ardour*[13] was used to render the tracks. Note that due to its geometric distribution (along a North-West/South-East line), the ETS data set required only 3 loudspeakers. The next sections explain the sonification methods and choices.

4.2 Sonification: General considerations and issues

As stated in Sect. 1, the most widely used sonification technique for seismic data is audification (few studies used other techniques, as e.g., Matsubara et al. [54] or McGee and Rogers [55]), as the *direct playback of data values by con-*

[10] Information at https://pnsn.org/seismograms.

[11] Freely available through an interactive web interface: https://www.pnsn.org/tremor.

[12] Details available at http://www.fdsn.org/networks/detail/Y4_2014/.

[13] Ardour is an open source digital audio workstation available at https://ardour.org/.

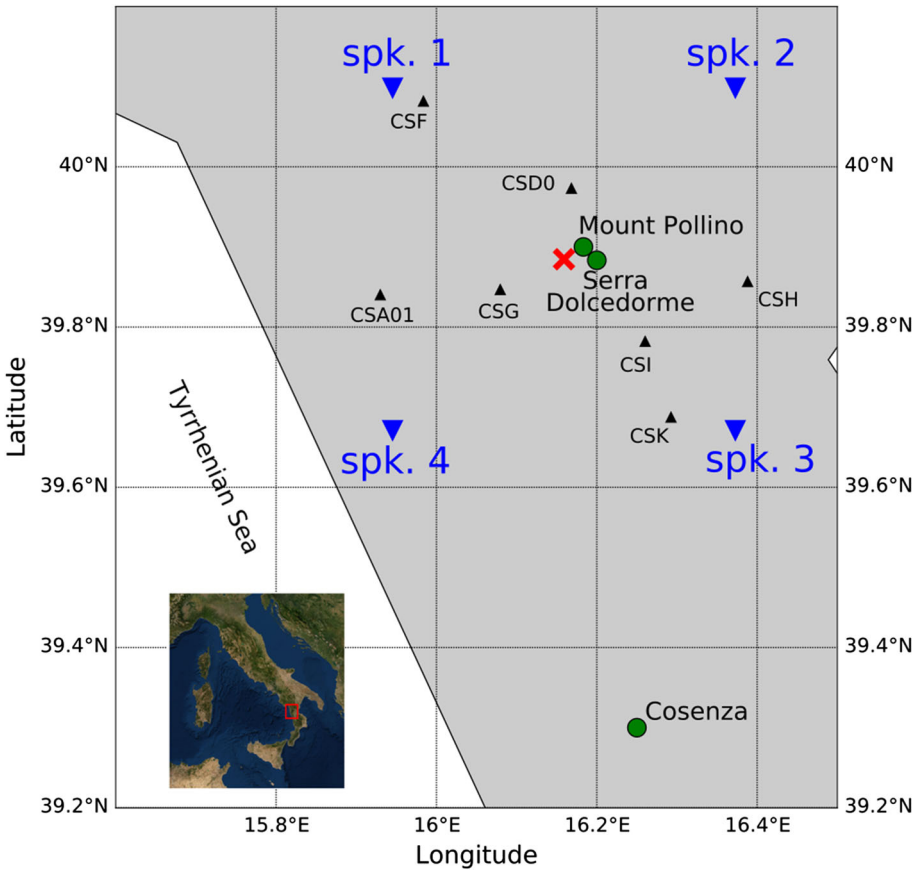

Fig. 3 Map view of the "Pollino" data set. The upward pointing triangles represent the selected seismic stations. Downward pointing triangles indicate the virtual position of the loudspeakers when playing the sounds (red cross indicates the center of the loudspeaker array). Green circles give landmarks to help the reader locate the area of interest: Pollino and Dolcedorme mountains, and the city of Cosenza (Calabria, Italy)

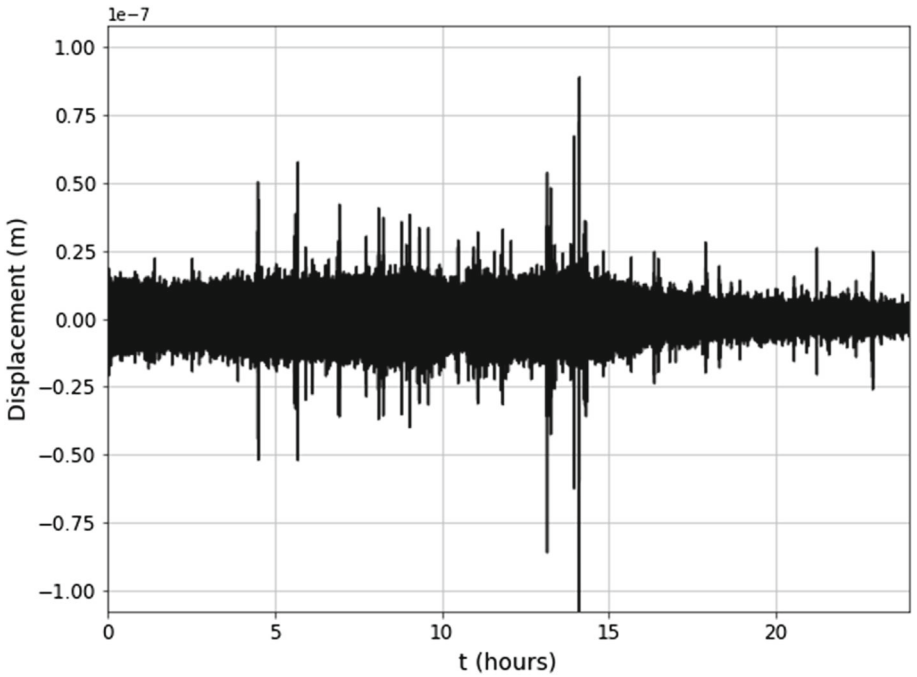

Fig. 4 Example waveform of a swarm recording in the Pollino region (Earth surface displacement data): 24-h activity on April 29th, 2015, recorded at station CSK

verting them into sound pressure [7, p. 288]. As explained in details in e.g., Dombois and Eckel [7] or Paté et al. [16], this technique consists in changing the time scale of the data with the effect of transposing the frequencies of the data into the human hearing range. In practice, it is enough to request from the computer a change in sampling frequency when playing back the sampled data points. A "speed factor", or "compression factor" can be defined as:

$$\mathscr{S} = \frac{F_{s,sound}}{F_{s,data}} = \frac{T_{data}}{T_{sound}}, \tag{1}$$

where $F_{s,data}$ is the data sampling frequency, $F_{s,sound}$ is the sound sampling frequency, T_{data} is the duration of the data signal, and T_{sound} is the duration of the resulting sound. Equation 1 shows that speeding up the playback of the data series (i.e. increasing the sampling frequency from $F_{s,data}$ to $F_{s,sound}$) is equivalent to compressing the signal duration (i.e., reducing the duration from T_{data} to T_{sound}). Another way to describe this speed factor is to think in terms of frequency shifting:

$$\mathscr{S} = \frac{f_{sound}}{f_{data}}, \tag{2}$$

where a specific frequency f_{data} from the spectrum of the data signal is shifted to a specific audio frequency f_{audio}. The speed factor can virtually take any value, resulting in some unusual values for $F_{s,sound}$: once scaled, the resulting sound data points might have to be resampled to a more classical audio sampling frequency (e.g., 44.1 kHz or 48 kHz).

Despite its simplicity, audification is not without problems. Of course it works better with data series that have a wave-like shape, and that contain a large number of points (otherwise the resulting sound would be too short to be relevant for listening purposes). It has been pointed out, e.g. by Dombois and Eckel [7], that compression should be used if the dynamic range of the data is larger than the practical dynamic range of our hearing (approximately 120 dB at 1000 Hz in an anechoic room, but closer to about 50 dB for complex signals in a real-life environment). In the present examples case, no compression was applied, and the downloaded data (quantized with at least 20 bits) was directly used, potentially preventing the listener from hearing the most quiet components.

It might also be—and this is true for many seismic recordings—that the frequency range of the data spectrum is so large that when it is shifted down to the audible range some side bands remain inaudible. In other words, the spectrum of seismic signals can be larger than the frequency range of human sensitivity. A choice then has to be made by filtering the data signal to focus on some frequency band only (depending on the phenomenon we want to observe: tidal waves, anthropogenic noise, micro-seismicity or larger earthquakes, …).

In the present case, the speed factors that were chosen brought the seismic frequency bands (1–10 Hz for Kilauea, 1–8 Hz for ETS, and 1–10 Hz for Pollino) to frequency ranges of fair to high sensitivity for the human auditory system (320–3200 Hz for Kilauea, 360–2800 Hz for ETS, and 1440–14400 Hz for Pollino). The ETS signals were even filtered to guarantee that the relevant frequency band was made audible (see Sect. 3.2).

As pointed out by Groß-Vogt et al. [56], another problem may occur with pure audification, that may require a *trade-off between the rhythmic structure and the displayed frequency range of individual events.* For the Pollino data set, the problem did not occur and pure audification was used. In the case of the Kilauea and ETS data sets however, the data signal duration (the relevant time window for each seismic event) and the data set time span (from first to last event in the data set) are so different that they prevent us from using only one speed factor. For illustration purposes, let's consider the extreme case where each signal is compressed to give 0.01-second long sounds, a duration so short that it is unlikely that listeners can proceed and interpret the features in the sound. In our case, pure audification with a single speed factor would result in speed factors of $\mathscr{S}_{kil} = \frac{80}{0.01} = 8000$ for Kilauea signals of original duration 80 s, and $\mathscr{S}_{ETS} = \frac{(6\times60)}{0.01} = 36,000$ for ETS signals of original duration 6 min. Given time spans of $T_{dataset,kil} = T_{dataset,ETS} = 13,219,200$ s (that is, 5 months) for the Kilauea and ETS data sets respectively, the whole soundtracks would last $T_{soundtrack,kil} = \frac{13,219,200}{\mathscr{S}_{kil}} = 1652.4$ s (a bit less than 28 min) for the Kilauea data set, and $T_{soundtrack,kil} = \frac{13,219,200}{\mathscr{S}_{ETS}} = 367.2$ s (a bit more than 6 min) for the ETS data set. These figures are much too long if the purpose of the sonification is to let the listener use her short-term memory to compare and recognize spatial and temporal patterns. Conversely, if it is sought to reach a 3-min long soundtrack (which already is quite long), it would result in speed factors of $\mathscr{S}_{kil} = \mathscr{S}_{ETS} = \frac{13,219,200}{3\times60} = 73,400$. And the resulting duration of each sonified event would be $T_{event,kil} = \frac{80}{\mathscr{S}_{kil}} \approx 1$ ms and $T_{event,kil} = \frac{(6\times60)}{\mathscr{S}_{kil}} \approx 5$ ms for Kilauea and ETS sonified signals respectively, that is to say too short sounds for the listeners to interpret them. Section 4.3 presents a possible solution to this compatibility issue between time scales.

4.3 Multiscale audification

The proposed choice here is to conduct audification with two concurrent speed factors. First, one speed factor, $\mathscr{S}_{track}$ is the speed factor of the track, that is the ratio between the time span of the observed phenomenon (5 months for Kilauea and ETS data sets, 24 hours for the Pollino data set) and the

soundtrack duration (100 seconds for Kilauea, 50 seconds for ETS, 60 seconds for Pollino; values chosen through informal tests among the authors as being correct tradeoffs between displaying enough details while keeping broader scale information perceptible, and minimizing cognitive load), such as:

$$\mathscr{S}_{track} = \frac{T_{dataset}}{T_{soundtrack}} \tag{3}$$

Resulting values for this first speed factor are $\mathscr{S}_{track} = 132,192$, $\mathscr{S}_{track} = 264,384$, and $\mathscr{S}_{track} = 1,440$ for Kilauea, ETS, and Pollino data sets respectively. s_{track} indicates the amount by which the duration of the observed set of events is reduced to give the rendered audio-visual track.

Then another speed factor $\mathscr{S}_{event}$ is introduced, as the speed factor at the event level, that is, the acceleration factor applied to the seismic signal to provide the sound signal during audification. It is the ratio between the duration T_{event} of a seismic event (the time window asked for download, centered around event time) and the duration T_{sound} of a single audio event such as:

$$\mathscr{S}_{event} = \frac{T_{event}}{T_{sound}} \tag{4}$$

For respectively the Kilauea and ETS data sets, seismic signals of duration $T_{event} = 80$ s and 360 s were used for producing single sounds of duration 0.3 s and 1 s, i.e. $\mathscr{S}_{event} = 267$, and 360, respectively. In the case of the Pollino data set, this second speed factor is not relevant and only $\mathscr{S}_{track}$ is used, for no event is detected nor used in the sonification (see Sect. 3.3).

In contrast with the volcanic earthquakes of the Kilauea dataset, the waveforms of tremor events in the ETS dataset are not impulsive, and when audified they would only sound like noise. In order to give each sound a dynamic inflection, each ETS waveform is multiplied waveform with a smooth envelope, composed of a short attack and a long release. We introduce here an envelope $e(t)$ defined by an upside down Lennard-Jones potential [57, p. 32, Eq. 1.3–27], a function that satisfies the shape that is sought:

$$e(t) = A\left(\frac{1}{(St+1)^{12}} - \frac{1}{(St+1)^{6}}\right), \tag{5}$$

where A scales the amplitude of the envelope, S controls the time scaling of the envelope. This envelope also acts as a taper (fade in and out) for the sounds, preventing clipping artifacts when the event sounds are summed together on the audiotrack.

In short, the basic audification method for the Kilauea and ETS soundtracks consists in placing each individual audified waveform (speed factor $\mathscr{S}_{event}$) at the correct relative time in the scaled timeline (speed factor $\mathscr{S}_{track}$). These baseline sonifications are called "K1" and "E1" (see Table 1 for a summary of the different sonification methods implemented). As so many events (more than 30, 000) are sonified in a so short period of time (less than 100 seconds), one might wonder whether the use of the audified, very complex waveform is not too subtle to be audible and whether it may be preferable to use simpler, more symbolic sounds. These alternative sonifications are presented in Sects. 4.4 and 4.5 present two methods for replacing the original waveform in the audification.

4.4 Alternative for audification 1: Envelope-modulated noise

The first alternative for audification is identical in principle to the multiscale audification of Sect. 4.3, the individual events being this time not represented by the audified original waveform, but by a simplified waveform. Each individual seismic waveform has its envelope extracted through a Hilbert transform. The envelope is then multiplied by a broadband, Gaussian noise, either low-pass (Butterworth filter of order 5, cut-off frequency at 300 Hz) or band-pass (Butterworth filter of order 5, cut-off frequencies at 100 and 3000 Hz) filtered. The idea is to keep the global temporal evolution of each seismic waveform unchanged, but to simplify all other aspects (in particular spectral aspects). This alternative method is called "envelope-modulated noise".

This first sonification alternative is applied to the ETS and Pollino data sets, under the code names "E2" and "P2", to which code "a" (low-pass filtering) or "b" (band-pass filtering) is appended (see Table 1).

4.5 Alternative for audification 2: Symbolic sounds

The second alternative is also identical in principle to the multiscale audification described above. The idea here is to represent events with even simpler sounds that elude both temporal and spectral features, only keeping the information on hypocenter location, onset time, and magnitude. Also, previous audifications (baseline and alternative 1) are very dense in terms of number of sounds per second, and the present alternative aims at using shorter sounds to obtain a more sparse soundtrack and help discriminating individual events that might otherwise superimpose. Consequently, individual events are this time replaced in the soundtrack by sonic icons of durations $t_{click} = 0.3$ s for Kilauea and $t_{click} = 0.15$ s for ETS (i.e., durations that are equal to or shorter than the audified waveforms with factor $\mathscr{S}_{event}$). Three versions of this sonic icon are proposed (see Fig. 5):

a a "click", i.e. a broadband noise (Gaussian distribution) with a temporal envelope $e(t)$ that exponentially goes up (10% of its duration) and then down (the remaining 90%),

Table 1 Summary of the sonification methods implemented, with their code names, corresponding data sets, and number of active loudspeaker in the proposed setup

Data set	Sonification method	Type of sound or filter	Code	Number of loudspeakers
Kilauea	Multiscale audification	Directly audified data	K1	4
	Symbolic sounds	Click	K3a	4
		Filtered click	K3b	4
		Desc. chirp	K3c	4
		Desc. chirp (start freq. $\propto$ magnitude)	K3d	4
ETS	Multiscale audification	Directly audified data	E1	3
	Envelope-modulated noise	Low-pass (300 Hz)	E2a	3
		Band-pass (100–3000 Hz)	E2b	3
	Symbolic sounds	Click	E3a	3
		Desc. chirp	E3c	3
Pollino	Multiscale audification	Directly audified data	P1	4
	Envelope-modulated noise	Low-pass (300 Hz)	P2a	4
		Band-pass (100–3000 Hz)	P2b	4

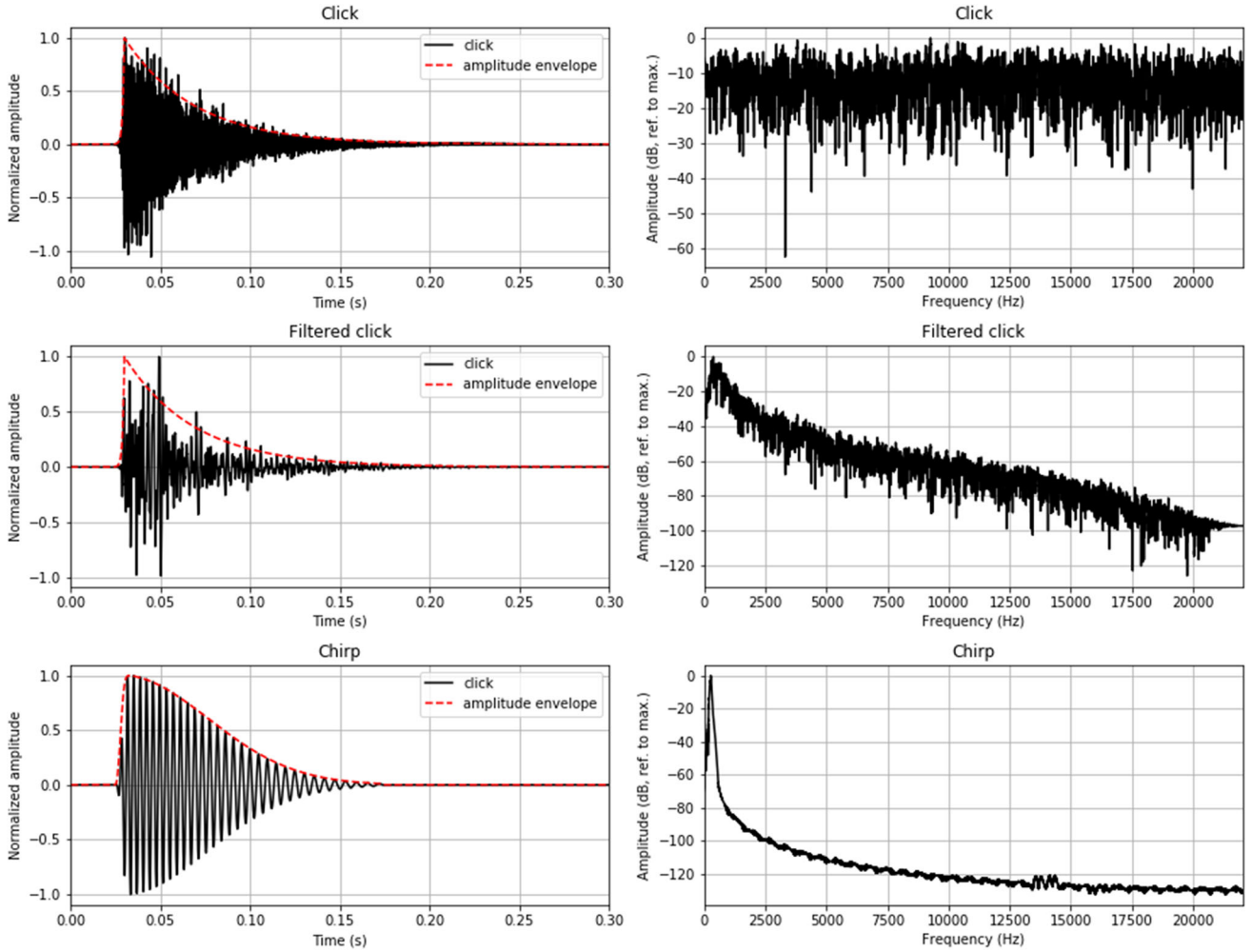

Fig. 5 Audification alternative 2. Three sonic icons proposed for use instead of the audified original seismic waveform. Left panels show the temporal forms and right panels the spectra of, from top to bottom, the click (symbolic sound "a"), the filtered click (symbolic sound "b"), the logarithmically descending chirp (symbolic sounds "c" and "d")

according to:

$$e(t) = \begin{cases} e^{\frac{\alpha_1}{\delta}(t-\delta)} & \text{if } 0 < t < \delta \\ e^{-\frac{\alpha_2}{t_{click}-\delta}(t-\delta)}, \end{cases} \tag{6}$$

with $\alpha_1 = 35$, $\alpha_2 = 7$, and $\delta = 0.03$ or 0.015 s (10% of the duration). These values result from an iterative design process by the authors.

b a "filtered click", i.e. the same click as above, this time filtered around 4, 410 Hz (Butterworth band-pass filtered of order 1 and quality factor 1).

c a logarithmically descending chirp from $f_{start} = 300$ to $f_{end} = 150$ Hz, multiplied by a skewed Gaussian defined as:

$$e(t) = \frac{1}{2} e^{-\left(\frac{t-\mu}{\sigma\sqrt{2}}\right)^2} \left(1 + erf\left(\frac{\alpha t}{2}\right)\right), \tag{7}$$

with $\sigma = 0.65$ (spread factor), $\alpha = 50$ (skewness factor), and erf the Gauss error function classically defined as $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\xi^2} d\xi$. These values result from an iterative design process by the authors.

d (for Kilauea data set only) the same logarithmically descending chirp, with f_{start} linearly mapped to the magnitude of the current event, and $f_{end} = \frac{f_{start}}{2}$. In order to obtain more bassy, deeper sounds for higher-magnitude events, the mapping between magnitude has an inverse polarity, so that the maximum magnitude ($M = 6.9$ for the Kilauea data set, only data set with magnitude information) in the data set is mapped to $f_{min} = 120$ Hz, and the minimum magnitude ($M = 0.5$ for the Kilauea data set) in the data set is mapped to $f_{max} = 600$ Hz.

This second sonification alternative, called "symbolic sounds", is applied to the Kilauea and ETS data sets, with code names starting with "K3" and "E3" and followed by the letter "a", "b", "c", or "d" for using respectively a click, a filtered click, a chirp, or a chirp whose start frequency depends on magnitude (see Table 1).

4.6 Spatialization

For all data sets, the "distance-based amplitude panning" (DBAP) method [58] is used for spatializing the sound. This method emulates the sound amplitude losses due to propagation in the air: The more distant a source is from the listener, the lower its amplitude at listening point. Similarly in the DBAP method, each sound is rendered by all loudspeakers, only the amplitude of each loudspeaker differs: The closer the source to a loudspeaker, the louder the loudspeaker will render the source's signal. Note that this method only implements the amplitude losses due to "geometric attenuation", i.e. as the wave travels further from the source, the initial energy is distributed over larger and larger surfaces, resulting in a decrease in intensity. The DBAP method, as it was first introduced, and as it is used here, does not implement other phenomena, e.g. the Stokes's law describing an attenuation that is frequency-dependent due to air viscosity effects.

In practice the amplitude a_k at which a source signal is rendered by loudspeaker k is computed via a 2-step procedure. First a normalization constant C is computed as:

$$C = \sqrt{\frac{1}{\sum_{k=1}^{K} \frac{1}{d_k^2}}}, \tag{8}$$

where K is the number of speakers (3 or 4 here). Then a_k is obtained as:

$$a_k = \frac{C}{d_k}, \tag{9}$$

where d_k is the distance between loudspeaker k and the source. Note that the loudspeaker positions are those of the "virtual" loudspeakers, as shown in Figs. 1, 2, and 3, i.e. their locations "as if" they stood within the geographical data space.

The Kilauea data set has a catalog that lists seismic events with their precise epicenter location, thus each individual sound (for each event) is spatialized, and the distance between speaker and source is in this case defined as the distance between the epicenter location of the current event and the virtual position of the loudspeaker.

For the ETS data set that has poorly located events, the sounds for each individual event are spatialized according to the distance between loudspeaker k and the location of the station that was selected for this event. Sameways for the Pollino data set that has no events: the audified sound tracks for each of the selected 7 stations are spatialized according to the distance between loudspeaker and stations.

At the end of the sound production procedure, and in order to avoid clipping, the tracks are normalized in amplitude with respect to the highest amplitude value encountered in all tracks. Then the tracks are converted to audio files.

5 Visual display

5.1 Visual projection setup

On each of both rendering sites, a video-projector of the brand *Optoma*, model *HD143X* was used. The videoprojector was put on a table at a height of approximately 0.8 m. The table was placed mid-way between loudspeakers 1 and 2 (see Figs. 1, 2 and 3), projecting on the next wall outside

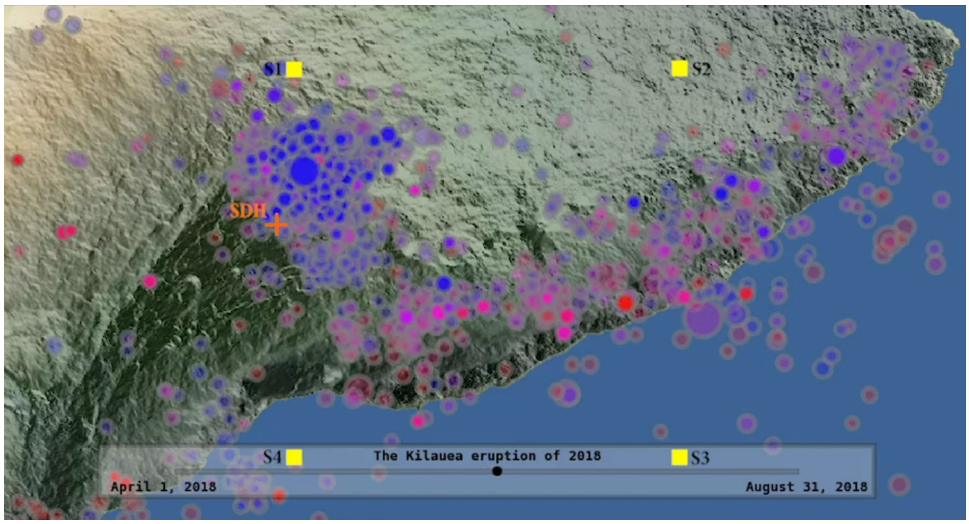

Fig. 6 A screenshot from the movie visualization of the Kilauea data set. It includes a progress bar showing the time line as well as colored dots popping at (scaled) onset time of events and location of epicenter. The color of the dots is according to the depth of the corresponding event, from 10 km below the surface (red) to the surface level (blue). The size of the dots varies with the magnitude of the corresponding events: the higher the magnitude, the larger the dots. Yellow squares indicate the virtual position of the loudspeakers, the attached mention reading "Si" stands for speaker number i ($= 1 \ldots 4$). The orange cross shows the selected seismic station, SDH

of the square defined by the loudspeakers, at a height convenient to the standing spectator. The dimensions (height × width) of the projected image were 0.9×1.7 m, 0.9×0.9 m, and 0.9×1.4 m for the Kilauea, ETS, and Pollino movies respectively.

5.2 Kilauea volcano

The visualization method for the Kilauea data set inherits from the "catalog movies" previously published by e.g., the "Seismic Sound Lab"[14]. A geographical map of the area of interest, with topological information, serves as the basis canvas. The animation consists in displaying colored dots popping at the time and location of the events in the catalog. The color and size of the dots respectively depend on the depth (hue-saturation-value color space from red to blue) and magnitude (size scales linearly with earthquake magnitude) of the corresponding events. The dots appear at the onset of each event and after 1 s their opacity begins to decay from 1 to 0.3 over a duration of 4 s. Also, markers show the virtual position of the 4 loudspeakers (yellow squares with speaker number attached). The sound and image tracks are synchronized, and a moving cursor at the bottom of the image indicates the current position in data time (i.e. from April 1st to August 31st, 2018). A screenshot of the movie rendering for the Kilauea data set is shown in Fig. 6.

[14] See http://www.seismicsoundlab.org/?page_id=157.

5.3 Episodic tremor and slip in Cascadia

The visualization for the ETS data set is inspired by the catalog movie of Sect. 5.2. A schematic map is displayed and remains for the whole duration of the movie, showing geographic (coasts, main cities) and geodynamic (Cascadia subduction trench) features, as well as the virtual position of the loudspeakers (white crosses at the corners of the screen). Visual events are superimposed onto this map, at the time of the event and identified position of its hypocenter. A counter at the bottom-left corner indicates the current position in data time (i.e., from July 1st to November 30th, 2012).

Even though discrete tremor source locations are identified in the dataset we use, they only represent the most probable source location for a segment of a continuous, diffuse tremor signal, emitted from a relatively broad source region in which an active tectonic process (slow-slip, fluid pressure diffusion) drives the seismic activity. The patterns of interest for us are thus drawn by the bulk time-space behavior of these sources. The visualization method presented here for tremor activity deliberately keeps and renders this "diffuse", "swarm-like" aspect of the data, in order to focus on the large scale patterns that the underlying driving tectonic process draws.

The event representation loosely corresponds to a 2D heat map of tremor activity in space. Brightest colors show the highest density of active tremor sources. For each event, a 2D Gaussian distribution centered on the epicentral location of the source is drawn. As they are stacked for each ani-

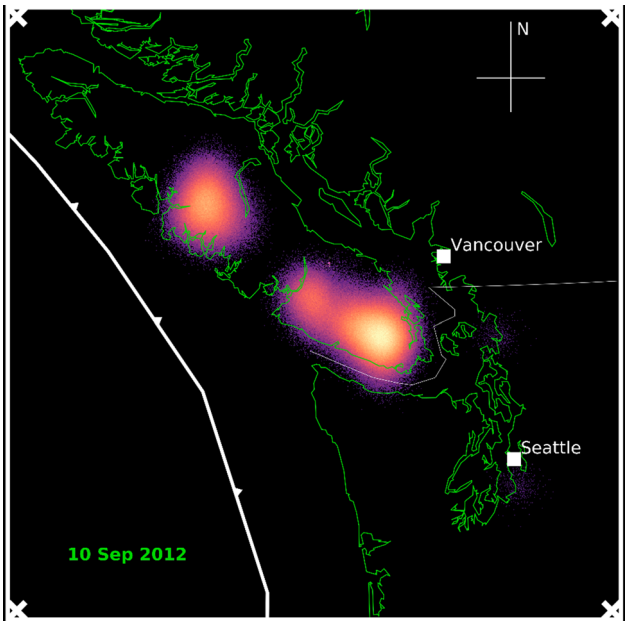

Fig. 7 A screenshot from the movie visualization of the ETS data set. Events are plotted at their (scaled) onset time and geographical position of epicenter as a colored haze extending in space. The progression of time is indicated as a date in the lower-left corner. The coast is represented with solid green lines, and the surface trace of the Cascadia subduction fault with a thick solid white line. The virtual position of loudspeakers is indicated with white crosses in the corners of the area. For guiding the viewers, a compass points towards the North, and the cities of Vancouver and Seattle are indicated

mation frame, it creates a haze that represents the diffuse activity zone, and more loosely, the location of the underlying tectonic process that drives the activity. A screenshot of the movie rendering for the ETS data set is shown in Fig. 7.

5.4 Pollino earthquake swarm

The Pollino data set consists in continuous data streams, and no identified events. A visualization similar to the two other data sets is therefore impossible. For this reason, the movie for the Pollino data set is very simplistic and certainly less informative than the previous ones. This lack of visualization clearly favors the use of audition for analyzing the data. This is very much in line with ongoing and former research in our group that have hypothesized that our auditory system can be used for data exploration purposes: once salient features are identified by the ears, the way is cleared for visualizing them.

The proposed visualization relies on the sliding bar paradigm (see a screenshot in Fig. 8). Two immobile panels present geographical maps, one at a global scale and national level (top left), the other one is more schematic and shows the position of the 7 selected stations and the virtual position of the loudspeaker (bottom left). From the 7 stations, 3 are further chosen, which span the whole geographical area of interest (CSD0 at the North, CSI at the South, CSA01 at the West). The right panels present the entire temporal traces from these three stations, as standard waveform plots. The color of the waveforms and of the stations on the geographical map are consistent. A moving vertical red bar shows to the listener the current playback position (in data time, that is over the 24-hour range). Also, a catalog of seismic events having occurred on the same day, with epicenters located nearby Italy was requested: their occurrence time is plotted with vertical black bars. These events have low magnitudes (1.5 to 3.6) and epicenters remote from our region of interest, hence they are not visible on the waveforms, but are intended to be guides to the ears. In fact, careful listening to the audifications sometimes reveal interesting sound changes at these very time positions…

6 Discussion

This article described methods to represent seismic data in the form of audio-visual clips that are dynamical i.e., that represent both spatial and temporal structures of the data. What motivated our work (see Sect. 2) is the observation that data visualizations rarely include sounds that are generated by the data, and data sonifications rarely include visuals that are generated by the data. An additional observation was that the use of space as a sound dimension is also quite rarely used in the field of sonification. When taken separately, the elements involved in our propositions were already known and used in the fields of data sonification and visualization: audification, spatialization, catalog movies, waveforms with moving bars, etc. In order to allow the visualization to represent patterns in time, and the sonification to represent patterns in space, the main contribution of the research presented in this article is (a) to bring all these elements together i.e., to include visuals, sounds, and spatialization in the same representational object; and (b) to adapt the representation methods to the specificities of each data set. This latter point led us to choose deliberately different data sets for testing the methods.

Two data sets (Kilauea and ETS) can be described using two different time scales: the time window of the data set, which is relevant for the identification of geophysical patterns vs. the single event time scale, which is relevant for the identification of local features. This motivated the development of a "multiscale audification", using different speed factors. Such a method allows the user to access and switch the focus between the analysis of single events, and the analysis of event clusters.

The sound tracks generated by multiscale audification however result in the juxtaposition of thousands of audified waveforms in a short time. These single sounds pass by so

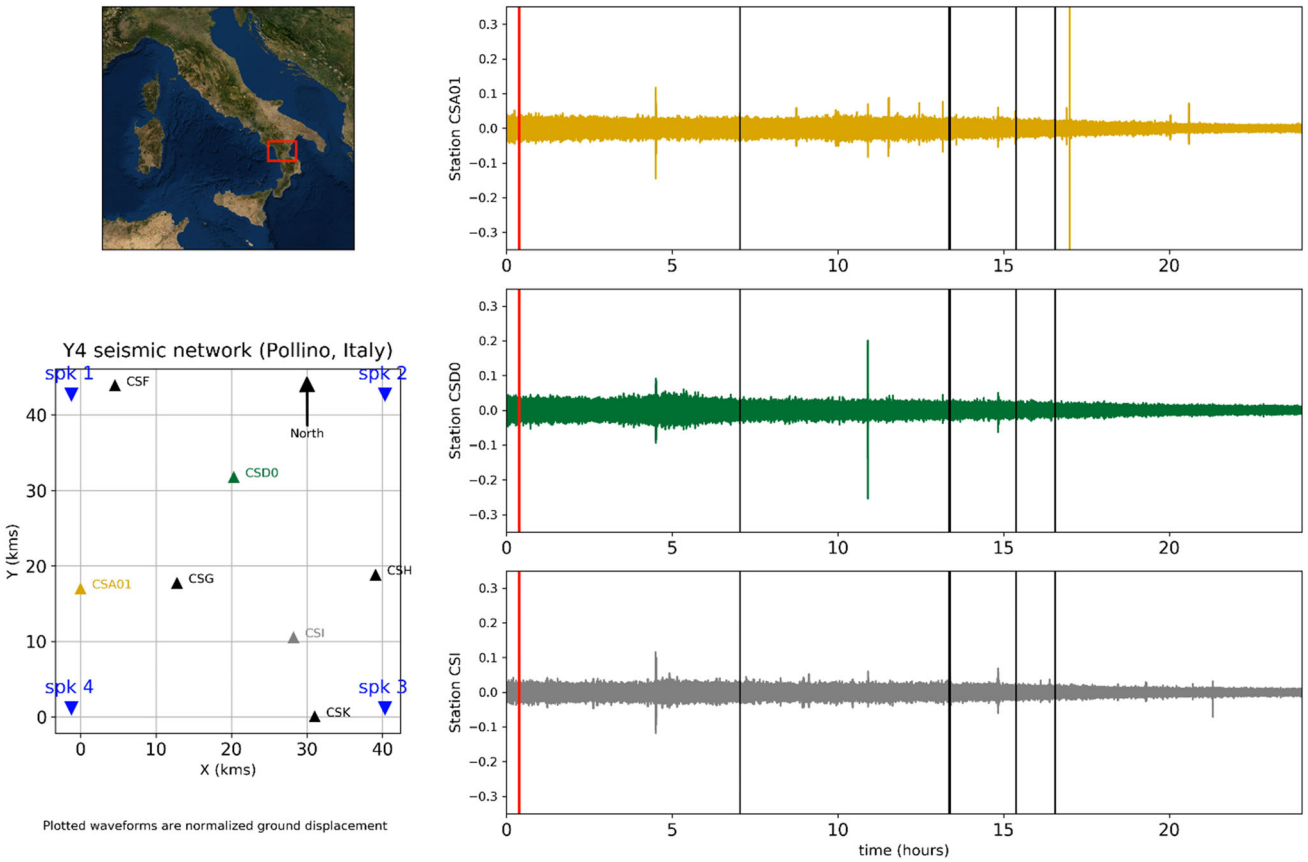

Fig. 8 A screenshot from the movie visualization of the Pollino data set. An inset in the upper-left corner represents a larger map to help the reader locate the area of interest. The lower-left corner shows a static map of stations (upward pointing triangles) and virtual position of loudspeakers (blue downward pointing triangles). The right part of the movie shows the recorded waveforms recorded at 3 stations (yellow, green and gray for stations CSA01, CSD0 and CSI), with a moving vertical red bar indicating the current (scaled) time position, and vertical black bars indicating some events detected outside of, but nearby the area of interest

quickly that they are hard to analyze separately, and their complexity increases the cognitive load of the listener too much. Thus the present article proposed to simplify the audification in substituting the audified waveforms with simpler, more symbolic sounds. Such "symbolic sounds" resemble "auditory icons" or "earcons" [37], and can be tuned according to the quantity of information the designer wants to convey through them. This article proposed the use of the following sonic objects, ordered here in increasing order of quantity of information given by the sound: simple or filtered clicks (indication of the presence/absence of an event), chirps with magnitude dependent-frequency (adding a symbolic representation of a valuable geophysical information), or "envelope-modulated noise" (adding magnitude information as sound amplitude, and representing the temporal evolution of the wave). We leave to future studies the definition of new symbolic sounds or the improvement of existing ones (e.g. filter according to the depth of the event), as well as the deeper investigation of the effects of combining audio and visuals [39].

Each data set has its specificities, so that the methods for representing them should be adapted. A "catalog movie" with individual, symbolic sounds is quite natural for the Kilauea data set, with events that are well defined in time and space. On the contrary, the tremors from the ETS data set resemble long-lasting and low-amplitude modifications of the background noise i.e., "events" that are loosely located in time and space: the migration of fuzzy clouds of points with sounds made of envelope-modulated noise is, to our mind, a better way to represent the "nature" of the data. Even further on the event/noise continuum, the earthquake swarms in the Pollino data set are more patterns in the background noise (that remain to be identified, maybe thanks to sonification, maybe not) than signals that stand out from the noise. In this case where the relevant data features are unknown, it makes more sense for now to proceed to "pure audification" (or envelope-modulated noise, to let the listeners focus on temporal features only), and stick to simplistic visuals.

This article focused on presenting the methods and it is behind its scope to conduct and analyze a structured percep-

Fig. 9 View of the 4-loudspeaker setup as installed in Marseille, France, during the demo session at CMMR 2019

tual experiment e.g., listening tests. However, we were able to gather a few comments during our two public presentations (see a picture of our setup at CMMR, Marseille in Fig. 9). Due to the format of these public representations (Open Lab Days in Lille, Demo during the CMMR conference in Marseille), including very limited time slots and user availability, as well as unsuitable conditions for perceptual testing, formal evaluations unfortunately couldn't be performed. The audience was quite diverse: university students, faculty and staff in Lille, but no experts in geosciences or in sound-related fields ; students and researchers in Marseille, with an expertise in sound-related topics (not in geosciences): computer science, sound design, acoustics, applied maths and signal processing, musicology, etc. The later audience provided more feedback, both in quantity and in quality, but the opinions of users from both audiences seemed to roughly agree.

In the case of the Kilauea data set, the potential benefit of combining sounds and visuals can be expressed as follows. First, the spatial migration (converging towards the crater) of events should already be clear enough in the movie, due to our well-developed spatial acuity in the visual domain. However, almost everyone reported to have heard this spatial migration: this certainly is the result of our the spatialization of sound, but the simultaneous and synchronous visual display may have helped a lot in this, by presenting the data in a consistent and congruent way. Also, this migration in space only makes sense because events move over time. The visuals account for the time evolution using colored dots popping at specific times and decaying at a certain rate: in the presence of many events/dots and at a short time scale, such time patterns might sometimes be hard to perceive, due to, e.g. retinal persistence. Due to the better time resolution of our auditory system compared to our visual system, the soundtrack made of sound events synchronized with the colored dots presumably helped perceive the temporal aspect of the data, therefore the motion of events, therefore their spatial migration.

Then, still considering the Kilauea data set, the temporal pattern of foreshocks sequence followed by main event (and no aftershock) was clearly perceived and could be analyzed as a rhythmic pattern: listeners clearly identified a periodic pulsation, with slight deviations in tempo that also could be distinguished, maybe calling for a geophysical interpretation. Replacing the audified events with chirps with magnitude-dependent frequency was even more informative to the listeners, as they could further parse the rhythmic patterns as clusters of *treble*[15] events (low magnitude) ending with a *bassy* event (high magnitude).

In the case of the ETS data set, the combination of sounds and visuals, and in particular the sound spatialization, allowed listeners to precisely track the tremor migration in the spatial audio. Again, the visualization of clouds of points moving in a way that is congruent to the spatial audio, may have increased the users' attention to the spatial patterns: some users explicitly acknowledged that the visuals were helpful to understand the sound. The simplification of the audified waveforms to envelope-modulated noise was not detrimental to the understanding of tremor migration, as one listener said: *[…] noise is less aggressive and still good for spatialization*.

The use of spatial audio for the Pollino data set drove the listeners to walk from loudspeaker to loudspeaker, trying to follow migrating phenomena, and to tell features common to several channels (presumably a meaningful signal) from features only present in a single channel (presumably noise in the sense of meaningless data, or artefacts due to the sensor). As described in Sect. 5.3, the visuals for the Pollino data set showed catalog events with vertical black bars (i.e. seismic events that are "large" enough, not considered as background

[15] Words uttered by the listeners are in italic font.

noise any longer). Some users tried to identify changes in sound at the very moments of these events. Interestingly, no regularity was found in this respect: some sound changes seem to be related to the occurrence of seismic events, and some other seismic events have no noticeable influence on the sound (one may hypothesize that the local background noise and swarm properties "mask" other greater but more distant events). An ongoing perceptual study in our team uses the Pollino data for the study of whether spatialization of audified signals (no visuals) helps understanding the migration of micro-seismic activity, whether the human auditory system is able to detect patterns in the background noise, and whether listeners agree on the specific patterns they detect, information that would provide much valuable insights for geoscientists.

In short, this article presented methods for the representation of seismic data, combining sonification, visualization, and spatial audio. The combined use of these display techniques, as well as their adaptation to the specificities of individual data sets (e.g., multiscale audification ; replacing simple audification by more simple, symbolic sounds ; hazy colored spots for visualizing tremors) is believed to provide the user with increased immersion and with more information without increasing her cognitive load (this latter point being supported by the fact that most users seem to have still been able to interpret the temporal and spatial aspects of the data even when the original waveforms were greatly simplified, as described in Sects. 4.4 and 4.5). Such enhanced data representation may prove helpful for analyzing and understanding data that is not yet fully described by current theories: this must be investigated by perceptual studies (Dubus and Bresin [59] alerted some years ago on the lack of user evaluation of auditory displays, and this observation may even more so apply for multimodal displays).

Data and codes

The webpage https://parthurp.github.io/homepage/SpatialSeismicSoundscapes_article2021.html links to the sounds and movies presented in this article. Provided are movies without sound, to be combined with the soundtracks available as separate tracks (i.e., one .wav file per channel). Also, movies with two-channel soundtrack were rendered for illustration purposes, ready to be played on more classical stereo setups. Note that the stereo soundtracks for these movies were generated using the DBAP method presented in this article, but with a different positioning of loudspeakers: left and right of the geographical area for Kilauea, upper-left and lower-right corners for ETS and Pollino. Besides, examples of the individual sonic icons are provided too: click, filtered click, descending chirp, straight audification.

Acknowledgements The authors would like to thank everyone who allowed them to make the audio-visual representations public: thanks to the CMMR organization team and attendance, thanks to the organizers and audience of the Open House days at LDEO/Columbia University, Junia, and Université Catholique de Lille, for giving helpful feedback. Thanks also to Josh Crozier from the Department of Earth Sciences at the University of Oregon for generating the Kilauea waveform dataset from the catalog.

Funding No funding was received to assist with the preparation of this manuscript.

Data availability all datasets used in this study are available on the web at the following addresses (see also the manuscript):
- https://www.usgs.gov/volcanoes/kilauea/monitoring
- https://www.usgs.gov/observatories/hawaiian-volcano-observatory
- https://pnsn.org/seismograms
- https://www.pnsn.org/tremor
- http://www.fdsn.org/networks/detail/Y4_2014.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

1. Kramer G, Walker B, Bonebright T, Cook P, Flowers JH, Miner N, Neuhoff J (1999) Sonification report: Status of the field and research agenda. Tech. rep, National Science Foundation
2. Latour B (1986) Visualization and cognition: drawing things together. Knowl Soc 6(6):1
3. Kramer G (1994) Auditory display: sonification, audification, and auditory interfaces. Avalon Publishing, New York
4. Hermann T, Hunt A, Neuhoff JG (2011) The sonification handbook. Logos Verlag, Berlin
5. Bernstein J, Schönlein C (1881) Telephonische Wahrnehmung der Schwankungen des Muskelstromes bei der Contraction. Sitzungsberichte der Naturforschenden Gesellschaft zu Halle, pp 18–27
6. Dombois F (2008) The muscle telephone. The undiscovered start of audification. In: Kursell J (ed) The 1870s. In: Sounds of science–Schall im labor. Max Planck Institute for the History of Science, Berlin, Germany, pp 41–45
7. Dombois F, Eckel G (2011) Audification. In: Hermann T, Hunt A, Neuhoff JG (eds) The sonification handbook, chap. 12(Logos Verlag, Berlin, Germany), pp 301–324
8. Speeth SD (1961) Seismometer sounds. J Acoust Soc Am 33(7):909. https://doi.org/10.1121/1.1908843
9. Frantti GE, Levereault LA (1965) Auditory discrimination of seismic signals from earthquakes and explosions. Bull Seismol Soc Am 55(1):1
10. Volmar A (2013) Listening to the cold war: the nuclear test ban negotiations, seismology, and psychoacoustics 1958–1963. Osiris 28(1):80. https://doi.org/10.1086/671364
11. Hayward C (1994) Listening to the Earth sing. In: Kramer G (ed) Auditory display: sonification, audification, and auditory interfaces, chap. 15. Addison-Wesley, Reading, pp 369–404
12. Dombois F (2001) Using audification in planetary seismology. In: Proceedings of the international conference on auditory display (ICAD)
13. Dombois F (2002) Auditory seismology—on free oscillations, focal mechanisms, explosions and synthetic seismograms. In:

Proceedings of the international conference on auditory display (ICAD)
14. Kilb DL, Peng Z, Simpson D, Michael AJ, Fisher M, Rohrlick D (2012) Listen, watch, learn: SeisSound video products. Seismol Res Lett 83(2):281. https://doi.org/10.1785/gssrl.83.2.281
15. Holtzman B, Candler J, Turk M, Peter D (2014) Seismic sound lab: sights, sounds and perception of the earth as an acoustic space. In: Aramaki M, Derrien O, Kronland-Martinet R, Ystad S (eds) Sound Music Motion. Springer International Publishing, New York, pp 161–174. https://doi.org/10.1007/978-3-319-12976-1_10
16. Paté A, Boschi L, Le Carrou JL, Holtzman B (2016) Categorization of seismic sources by auditory display. Int J Human Comput Stud 85:57. https://doi.org/10.1016/j.ijhcs.2015.08.002
17. Paté A, Boschi L, Dubois D, Le Carrou JL, Holtzman B (2017) Auditory display of seismic data: on the use of expert's categorizations and verbal descriptions as heuristics for geoscience. J Acoust Soc Am 10(1121/1):4978441
18. Boschi L, Delcor L, Le Carrou JL, Fritz C, Paté A, Holtzman B (2017) On the perception of audified seismograms. Seismol Res Lett 88(5):1279. https://doi.org/10.1785/0220170077
19. Gaver WW (1989) The SonicFinder: an interface that uses auditory icons. Human-Comput Interact 4:67. https://doi.org/10.1207/s15327051hci0401_3
20. Neuhoff JG (2011) Perception, cognition and action in auditory displays. In: Hermann T, Hunt A, Neuhoff JG (eds) The sonification handbook, chap. 4. Logos Verlag, Berlin, pp 63–86
21. Walker A, Brewster S, Mcgookin D, Ng A (2001) Diary in the sky: a spatial audio display for a mobile calendar. In: Proceedings of the 5th annual conference of the British HCI Group (Lille, France), pp 531–539. https://doi.org/10.1007/978-1-4471-0353-0_33
22. Howard IP, Templeton WB (1966) Human spatial orientation. Wiley, New York no DOI
23. Deutsch D (2013) Grouping mechanisms in music. In: Deutsch D (ed) The psychology of music, chap. 6. Elsevier, Amsterdam. https://doi.org/10.1016/C2009-0-62532-0. 3rd edition
24. Bregman AS (1994) Auditory scene analysis: the perceptual organization of sound. MIT Press, Cambridge no DOI
25. Féron F, Frissen I, Boissinot J, Guastavino C (2010) Upper limits of auditory rotational motion perception. J Acoust Soc Am 128(6):3703. https://doi.org/10.1121/1.3502456
26. Blauert J (1983) Spatial hearing. MIT Press, Cambridge
27. Middlebrooks JC, Green DM (1991) Sound localization by human listeners. Annu Rev Psychol 42(1):135. https://doi.org/10.1146/annurev.ps.42.020191.001031
28. Dunai L, Lengua I, Peris-Fajarnés G, Brusola F (2015) Virtual sound localization by blind people. Arch Acoust 40(4):561. https://doi.org/10.1515/aoa-2015-0055
29. Guillaume A, Rivenez M, Andéol G, Pellieux L (2007) Perception of urgency and spatialization of auditory alarms. In: Proceedings of the 13th international conference on auditory display (ICAD), Montréal, Canada
30. Parente P, Bishop G (2003) BATS: The Blind Audio Tactile Mapping System. In: Proceedings of ACM south eastern conference
31. Loeliger E, Stockman T (2012) Wayfinding without visual cues: evaluation of an interactive audio map system. Interact Comput 26(5):403. https://doi.org/10.1093/iwc/iwt042
32. Brungart DS, Simpson BD (2008) Design, validation, and in-flight evaluation of an auditory attitude indicator based on pilot-selected music. In: Proceedings of the 14th international conference on auditory display (ICAD), Paris, France
33. Kiefer P (ed) (2010) Klangräume der Kunst. Kehrer Verlag, Berlin
34. Brown LM, Brewster SA, Ramloll R, Burton M, Riedel B (2003) Design guidelines for audio presentation of graphs and tables. In: Proceedings of the international conference on auditory display (ICAD), Boston, MA
35. Roginska A, Childs E, Johnson MK (2006) Monitoring real-time data: a sonification approach. In: Proceedings of the 12th international conference on auditory display (ICAD), London, UK
36. Bonebright TL, Nees MA, Connerley TT, McCain GR (2001) Testing the effectiveness of sonified graphs for education: a programmatic research project. In: Proceedings of the international conference on auditory display (ICAD), Espoo, Finland
37. McGookin D, Brewster S (2011) Earcons. In: Hermann T, Hunt A, Neuhoff JG (eds) The sonification handbook, chap. 14. Logos Verlag, Berlin, pp 339–362
38. Holtzman B, Candler J, Repetto D, Pratt M, Paté A, Turk M, Gualtieri L, Peter D, Trakinski V, Ebel D, Gossmann J, Lem N (2017) SeismoDome: sonic and visual representation of earthquakes and seismic waves in the planetarium. In: AGU fall meeting, New Orleans, LA
39. Barth A, Karlstrom L, Holtzman B, Niyak A, Paté A (2020) Sonification and animation of multivariate data illuminates geyser eruption dynamics. Comput Music J (submitted)
40. Peng Z, Aiken C, Kilb D, Shelly DR, Enescu B (2012) Listening to the 2011 Magnitude 9.0 Tohoku-Oki, Japan. Earthq Seismol Res Lett 83(2):287. https://doi.org/10.1785/gssrl.83.2.287
41. Dombois F, Brodwolf O, Friedli O, Rennert I, Koenig T (2008) Sonifyer—a concept, a software, a platform. In: Proceedings of the international conference on auditory display (ICAD)
42. Spence C (2007) Audiovisual multisensory integration. Acoust Sci Technol 28(2):61. https://doi.org/10.1250/ast.28.61
43. Hendrix C, Barfield W (1996) The sense of presence within auditory virtual environments. Presence 5(3):290. https://doi.org/10.1162/pres.1996.5.3.290
44. Viaud-Delmon I, Warusfel O, Seguelas A, Rio E, Jouvent R (2006) High sensitivity to multisensory conflicts in agoraphobia exhibited by virtual reality. Eur Psychiat 21(7):501. https://doi.org/10.1016/j.eurpsy.2004.10.004
45. Chouet BA, Matoza RS (2013) A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption. J Volcanol Geoth Res 252:108. https://doi.org/10.1016/j.jvolgeores.2012.11.013
46. Liang C, Crozier J, L K, M DE (2020) Magma oscillations in a conduit-reservoir system, application to very long period (VLP) seismicity at basaltic volcanoes: 2. Data inversion and interpretation at Klauea Volcano. J Geophys Res Solid Earth https://doi.org/10.1029/2019JB017456
47. Wu SM, Lin FC, Farrell J, Shiro B, Karlstrom L, Okubo P, Koper K (2020) Spatiotemporal seismic structure variations associated with the 2018 Klauea eruption based on temporary dense geophone arrays. Geophys Res Lett. https://doi.org/10.1029/2019GL086668
48. White RS, Drew J, Martens HR, Key J, Soosalu H, Jakobsdóttir SS (2011) Dynamics of dyke intrusion in the mid-crust of Iceland. Earth Planet Sci Lett 304(3–4):300. https://doi.org/10.1016/j.epsl.2011.02.038
49. Woods J, Winder T, White RS, Brandsdóttir B (2019) Evolution of a lateral dike intrusion revealed by relatively-relocated dike-induced earthquakes: The 2014–15 Bárðarbunga-Holuhraun rifting event, Iceland. Earth Planet Sci Lett 506:53. https://doi.org/10.1016/j.epsl.2018.10.032
50. Rogers G, Dragert H (2003) Episodic tremor and slip on the Cascadia subduction zone: the chatter of silent slip. Science 300(5627):1942. https://doi.org/10.1126/science.1084783
51. Frank WB, Shapiro NM, Husker AL, Kostoglodov V, Bhat HS, Campillo M (2015) Along-fault pore-pressure evolution during a slow-slip event in Guerrero, Mexico. Earth Planet Sci Lett 413:135. https://doi.org/10.1016/j.epsl.2014.12.051
52. Wech AG, Creager KC (2008) Automated detection and location of Cascadia tremor. Geophys Res Lett. https://doi.org/10.1029/2008GL035458

53. Roessler D, Passarelli L, Govoni A, Bautz R, Dahm T, Maccaferri F, Rivalta E, Schierjott J, Woith H (2014) Extended Pollino Seismic Experiment, 2014-2015, GFZ Potsdam (FEFI, Pompei, NERA projects). GFZ Data Services. https://doi.org/10.14470/L9180569
54. Matsubara M, Morimoto Y, Uchide T (2016) Collaborative study of interactive seismic array sonification for data exploration and public outreach activities. In Proceedings of ISon 2016, 5th interactive sonification workshop, Bielefeld, Germany
55. McGee R, Rogers D (2016) Musification of seismic data. In: Proceedings of the international conference on auditory display (ICAD), Canberra, Australia
56. Groß-Vogt K, Frank M, Höldrich R (2019) Focused Audification and the optimization of its parameters. J Multimodal User Interf. https://doi.org/10.1007/s12193-019-00317-8
57. Hirschfelder JO, Curtiss CF, Bird RB (1964) Molecular theory of gases and liquids. Wiley, New York
58. Lossius T, Pascal PB, de la Hogue T (2009) DBAP–distance-based amplitude panning. In: Proceedings of the international computer music conference (ICMC)
59. Dubus G, Bresin R (2013) A systematic review of mapping strategies for the sonification of physical quantities. PLoS ONE 8(12):e82491. https://doi.org/10.1371/journal.pone.0082491