Movement interaction with a loudspeaker: an index of possibilities

Georgios Marentakis
Department of Computer Science and Communication
Østfold University College
Halden, Norway
georgios.marentakis@hiof.no

ABSTRACT

Creating sound by manipulating the location of sound-producing objects using gestures is an interesting interaction paradigm. To understand it better, we analyzed videos of users interacting with 'Random Access Lattice'. In this sound installation, users move a loudspeaker to explore sound laid out in space using a time to space mapping. We performed an inductive analysis of user movement in relation to the visitor intention and the sonic outcome. We identified several body, hand, and grip gestures which were performed with different movement qualities to manipulate the loudspeaker at variable speeds, orientations, and body areas. These were used to search and trace the sonic material in a goal-oriented fashion but also to interact creatively with sound by looping and modulating sounds. We provide a visual index of our findings which can be used when designing gesture interactions with sound.

Authors Keywords

movement and gesture; sonic interaction design

C&C '22, June 20–23, 2022, Venice, Italy

ACM ISBN 978-1-4503-9327-0/22/06.
https://doi.org/10.1145/3527927.3533026

INTRODUCTION

Designing for gesture, movement, and the body is receiving significant attention in interaction design. Starting from manipulative and symbolic gestures in tangible [30] and mobile computing [45, 61] researchers look now into whole-body interaction and designing for movement. Gesture and movement-based interaction attributes increased importance to bodily, social, and contextual aspects as advocated by embodied interaction design [14].

Whole-body interactions originated in games and interactive virtual environments in which context movement was considered beneficial for immersion [49]. Eventually, movement-based interaction that brings the body to the fore of the design process emerged as a research objective [1, 38]. The relationship between gesture, movement, and sound is fundamental. Gesture and whole-body interactions are often combined with audio as output modality. This practice has roots in music and installation art [60]. More recent work investigates body centered auditory feedback [58], sonic interaction design [21], and gestures and audio for mobile interaction [8, 40, 51, 62].

The work we present here investigates a novel interaction paradigm: using gestures and movement to explore sound that has been laid out in 3D space. The interactive sound installation 'Random Access Lattice' by Gerhard Eckel is used as a departure point for this study as it introduces and realizes this novel interaction. An inductive analysis of the participant's movements as these were captured in video recordings during evaluation is used to illustrate the design knowledge embodied in the installation. These are visualized using images and movement sequences. Our intention is that the illustrated knowledge acts in a generative way to stimulate interaction design within this area. First, gesture and movement-based interaction with sound is reviewed. Subsequently, the installation, the methodology, and the results are presented. Finally, how the results illuminate gesture interaction with sound is discussed.

GESTURE AND MOVEMENT-BASED INTERACTION WITH SOUND

Gesture and movement-based interaction with sound has roots in computer music and interactive sound art [65]. Mid-air gestures in computer music appear in the Theremin and other similar instruments such as the Ondes Martenot or the Trautonium and the Lady's glove. In more recent works, gestures have been used for controlling sound synthesis models [63] but also sound spatialization [41]. Several tangible interfaces employ manipulative gestures for sound and music. A well-known example is the Reactable [31] or the Audiocubes [54]. The body has received significant interest in computer music with interfaces building on EEG but also EMG data [19, 60], providing biofeedback [4, 5], and for sonic interaction design [21].

Gestures are quintessential for sound production and their relationship to sound has been studied in detail. When asked to freely produce gestures in response to sound people typically perform the sound producing action and if not available focus on mimicking a prominent sound feature e.g. the evolution of the envelope [11, 24]. Such gestural sonic affordances have found their way in music and sound design applications [2, 12, 59].

Beyond musical applications, gesture interfaces to sound have been designed for interacting with speech messages such as voice notes [56], auditory menus [8, 40], memories [13, 50], audiography [15], checking the status of appliances [47], supporting vulnerable groups such as people with dementia, or for creating interactive furniture e.g., [3]. Wearable interfaces have also been designed, a few of them focusing on sound-producing fabrics [57]. Gestures play an important role for interactive sonification [25].

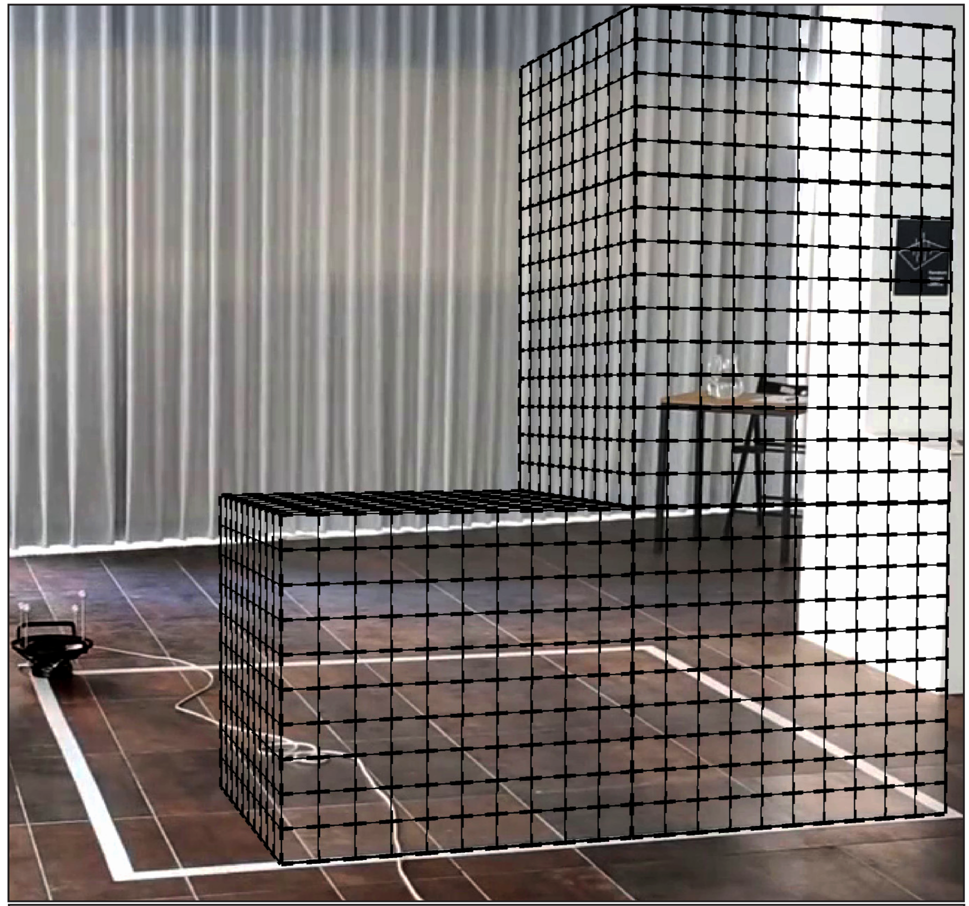
RANDOM ACCESS LATTICE

The research on movement interaction with sound presented here is based on the interactive sound installation 'Random Access Lattice' which was conceived and realized by Gerhard Eckel [16]. The installation operates using a loudspeaker on which reflective markers are attached so that its location can be tracked by an optical tracking system. The loudspeaker is silent when still and plays back audio from a collection of poems recited in 40 different languages when moved within a specified area. These are arranged in a three-dimensional grid, a lattice, forming a cube of 2 by 2 by 2 meters. Grid lines are parallel to one of axes of the surrounding cube. The grid is rather tight and has a large number of subdivisions. There is a different voice every 8 cm along each of the three axes in the cube. Along the two meters of the cube's lateral length seven seconds of sound are stored. Only if the loudspeaker is moved exactly along one grid line, the individual track stored there will be heard. Otherwise, the sound of neighboring lines is mixed to create a sonic outcome that varies depending on loudspeaker speed and exact location. This design results in that moving back and forth creates loops, whose sonic variability is increased when deviating from grid lines. Faster movements result in louder sound. The pitch of the sounds does not change with the speed of movement. This is achieved by using granular synthesis [16, 22]. Different tracks are engrained on the lattice for each of the principal movement directions. As a result if the direction of movement is reversed, another sound is heard and not the same sound played backwards. A rectangle on the floor marks the bottom side of the cube for staging.

RAL in Context

RAL is an example of tangible and embodied interaction with sound [16] and originates in sonic interaction design [21]. Quite striking and novel in RAL is the temporal and spatial coincidence of output sound and input gesture which are unified on the tracked loudspeaker. Most commonly in sonic interactions in the literature, the locations of the sound producing gesture and the sound output do not coincide as sound originates from loudspeakers or headphones away from the input action location e.g., in [34, 51]. Approximate spatial coincidence between user action and sound output occurs when interacting with smart sonic objects and loudspeakers [43], smart headphones [9] or mobile phones [64]. However, interactions with such devices not do involve translation in space and are mostly based on physical controls or touch gestures.

The loudspeaker in RAL can also be used as an instrument in a performative way. Gerhard Eckel refers to the tape bow violin by Laurie Anderson [35] in this context.



A staging of the installation showing the tracked loudspeaker and the marked installation operation area. A part of the lattice used to layout sounds in the installation, which extended over the whole area, is shown scaled in proportion.

Furthermore, he links Random Access Lattice to Random Access by Nam June Paik [48], and provides a possible interpretation of the moving loudspeaker as the head of tape player [16]. In Random Access, the visitor can use a detached tape recorder head to interactively run through audio tapes glued to the wall. This makes it possible to vary the sound outcome according to the location and speed with which users interact and refers to the effects applied on processed pre-recorded fragments, loops, and sampled sounds in tape music. Tape music is an integral part of musique concrète and the work of composers such as Pierre Schaeffer, Pierre Henry, Edgar Varese and John Cage. The complex relationship between the loudspeaker, the signal processing and amplification chain, and the room and the spatial separation between the location of the performer and the sound is characteristic of electroacoustic music [46], a point that is dealt explicitly by the design of RAL.

A further novel and exciting point in RAL is the spatial arrangement of the sound in 3D space using a time to space mapping. Time to space mappings are central to audio technology. They appear in early inventions, such as the phonograph, the vinyl plate, and the audio tape, and are also central to virtually all manipulations performed in contemporary Digital Audio Workstations. Such mappings have also been used in designing spatial audio interfaces. Kobayashi et al. [32] developed an audio browsing tool in which audio documents are placed in a circular auditory orbit around a user and read by orbiting speakers. Sawhney and Schmandt [53] map the time of arrival of events was mapped to different positions in circular auditory space.

METHODOLOGY

As mentioned above, RAL is unique both in that it unifies gesture and sound location and in the way with which it applies a time to 3D space mapping to engrave sound in space and support interaction. Consequently, RAL affords interactions that have not been observed before. In this sense, despite being routed in artistic research, RAL embodies knowledge related to designing movement interaction with sound. However, this design knowledge is only made visible in the participant interactions and the bodily ways they engaged with the installation. The research we present here is motivated by the need to uncover the movement design knowledge embodied in RAL. To answer our research question, we observe and document how users engage bodily with the installation during evaluation.

Evaluation methods from human computer interaction have been commonly applied to interactive art to resolve usability issues but also to provide insight into the visitor experience and contrast it to the artistic intention [7, 10, 29]. The applicability of evaluation methods in interactive art has been discussed widely, given that interactive art has different goals than human computer interaction. The primary concern of the artist is not to create a usable system nor a common experience. Consequently, usability and user experience evaluation methods do not have a summative but rather a formative function in the research process.

Close-up of the loudspeaker showing the handles used to manipulate it, the markers used to track its location, and the cable connecting to the sound system

Our methodology is based on the premise that evaluation can act as a tool to unlock, document, and present the design knowledge obtained as a result of the artistic research process. An interactive artwork embodies knowledge as the result of an artistic research process. Aesthetic knowledge derived from an aesthetic experience manifests itself at the sensory, emotional, and reflective level and cannot be fully expressed verbally. It may be argued that artistic research shares several aspects with design research. Both share roots in phenomenology and aim to generate knowledge albeit with a different focus. Artistic research is primarily concerned with aesthetic reflection and knowledge. Design research is considered to be the source of intermediate-level design knowledge that can be communicated using annotated portfolios [23], strong concepts [28], or experiential qualities [39]. In approaching RAL from a (movement) design research perspective methods from movement-based interaction become relevant.

Designing movement-based interaction typically involves observation coupled with movement design and enactment activities [18]. Observation is done using first, second, and third person methods. First-person methods sensitize designers and expose them to the experience of movement often under the guidance of a somatic connoisseur [55]. Second person methods allow designers to empathize with how movement is experienced by other persons. Third-person methods aim to objectify movement through observation and ethnography but also provide visual analysis and representations as when using movement sequences, Laban movement analysis, or spatial movement schemas [38]. Several systems automatically recognize Laban movement qualities so that these are used as an input method in interaction [17, 44].

Observing movement may be done before, during, or after the design process is finished. For example, in preparation for designing interaction [37] charts the movement of museum visitors. Several researchers sensitize themselves to movement in advance using movement awareness techniques or design and observe movement simultaneously [27, 33, 55]. For works dealing with novel interaction paradigms, observing and documenting movement in retrospect can provide material that operates in a generative way to stimulate interaction design in future scenarios. Loke et al [36] investigate movements performed when interacting with finished video games. This type of assessment-oriented knowledge is related to experiential qualities of the interaction developed here by the artist and can inform the design of interactions that fall within the same genre [26].

A similar approach is taken here which aims to complement the RAL experiment by bringing forward the (movement) design knowledge that emerged in this artistic research process. We focus on which gestures occur during interaction and which goals these gestures accomplish during installation evaluation using a third-person perspective.

Laban analysis is commonly applied when analyzing, referring or describing movement-based interaction. Laban analysis can be used to describe gestures using Labanotation. Labanotation is not used here for two reasons. The first is that although incorporating objects or props is possible in Labanotation, the objects typically handled are related to dance and ballet performances and are quite different from the loudspeaker we have here. The second reason is that Labanotation is not as widespread within interaction design and representations are not as informative for several readers. Movement sequences or silhouettes are used instead to describe the form of a gesture.

However, Laban analysis is important because it can describe differences in the performing qualities of gestures that have the same form. This is done using the Laban Effort analysis. This classifies gestures in terms of four dimensions: Space (direct/indirect), Time (quick/sustained), Weight (Light/Strong), and Flow (Free/Bound). The combinations of Space, Time, Weight, and Flow lead to definition of eight Laban Basic Efforts: Wring, Press, Flick, Dab, Glide, Float, Punch, and Slash which have been used extensively in the HCI literature [6, 17, 20, 36]. We make use of the Effort space and basic efforts in order to differentiate between gestures of similar form that are associated with different user intentions.

Data Collection

RAL has been evaluated as part of a bigger project looking at the evaluation of several sound installations using different evaluation techniques [42, 52]. Evaluation was not based solely on video observation but also collected other types of data and materials. The material we analyzed corresponds to video recordings of the beginning of evaluation sessions in which participants explored RAL at their own pace. The focus of this article is on bodily and movement aspects as these were derived in the analysis of the video recordings.

Data collection was performed on two occasions during installation evaluation sessions. Videos were analyzed in an inductive way and the gestures used by the users but also the tasks they served were coded and categorized focusing on the sonic output of the interactions. A brief presentation of the installation was given to participants before they started. An assistant coordinated video recording and helped with questions or difficulties. In total, there were sixteen participants, eleven participants that took part in the first and another five in the second evaluation session. Informed consent was obtained before evaluation started. The total duration of the videos was 274 minutes and included about 6000 coded segments. The corresponding videos were segmented, coded, and categorized. Segmentation was performed by an assistant and was verified with minor corrections by the author who then worked on developing the coding scheme presented next.

CODING SCHEME

The following gesture categories were identified when analyzing the dataset:

- Body gesture: refers to the body gesture performed while interacting and includes static (stand, sit), dynamic (turn, jump, step, and balance), and mixed gestures performed both as static and as dynamic (kneeing and bending).
- Hand gesture: refers to the way the loudspeaker was manipulated by the user's arm and hand gestures. Most gestures were dynamic (rotate, displace, swing, tilt, slide, shake, muffle) but there was also one static (hold) gesture.
- Grip gesture: refers to the grip used to hold the loudspeaker using the hands (single hand, both hands, or by using the cable)
- Gesture speed: refers to the speed with which the sound producing gesture was performed (low, regular, high, or mixed)
- Loudspeaker orientation: refers to the orientation of the loudspeaker cone relative to the floor and the user (upwards, downwards, inwards, outwards)
- Hand gesture location: refers to the area in which the loudspeaker moved relative to the body of the user (head, torso, or legs, floor, on or above user's head)

Most common body gestures were standing and stepping, while sitting and jumping were the least common. Most common hand gestures were displace and hold, while shake and muffle were least common. Most gestures were performed with both hands, and while almost all gestures were also performed with a single hand this happened less often. The cable was only used in combination with sliding (pulling) the speaker on the floor. We observed most of the possible body, hand, and grip gesture combinations in the dataset and not all combinations are possible to visualize here. All users had a 'regular' speed which they used for most of the gestures they performed and sped up or slowed down occasionally. The loudspeaker was most often held facing upwards and while all other orientations were observed they were not as common. Finally, most gestures took place in the user's torso area. The leg and head area followed while the floor and the area on the top or above the user's head were not used as often.

PERFORMING MODES

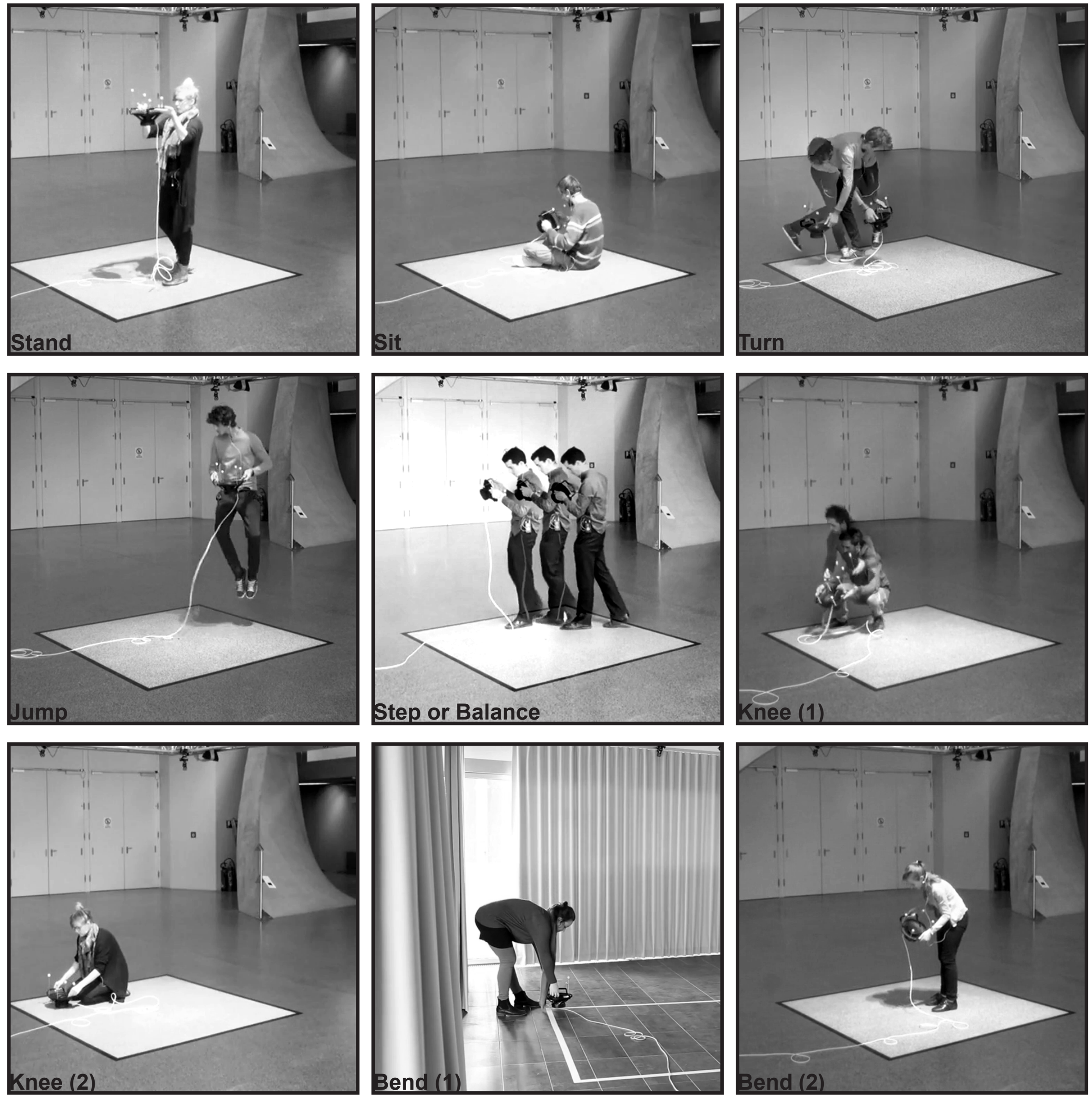
Body and hand gestures interacted significantly in the sense that some combinations were more common than others. Three most common clusters were identified. The first is associated with the displacing gesture which is most often performed in the standing position (stand & displace), less frequently while stepping, bending, or kneeing, and seldom while sitting, turning, balancing, or jumping. The second is associated with tilting and is most often performed when standing or kneeing (stand & tilt and knee & tilt), and less frequently while stepping, and very seldom while sitting, turning, balancing, or jumping. The third is associated with the holding gesture is most often performed while stepping (step & hold), less frequently while turning, balancing, bending, or kneeing, and seldom while jumping.

The gestures identified above were operationalized while exploring, interacting, and eventually performing the installation. This was done in four major ways (or modes) which were classified as: crossing, tracing, looping, and modulating. The modes are defined with respect to the sound that is produced as a result of the interaction. In many ways, the performing modes are closely associated with or directly describe the user intention.

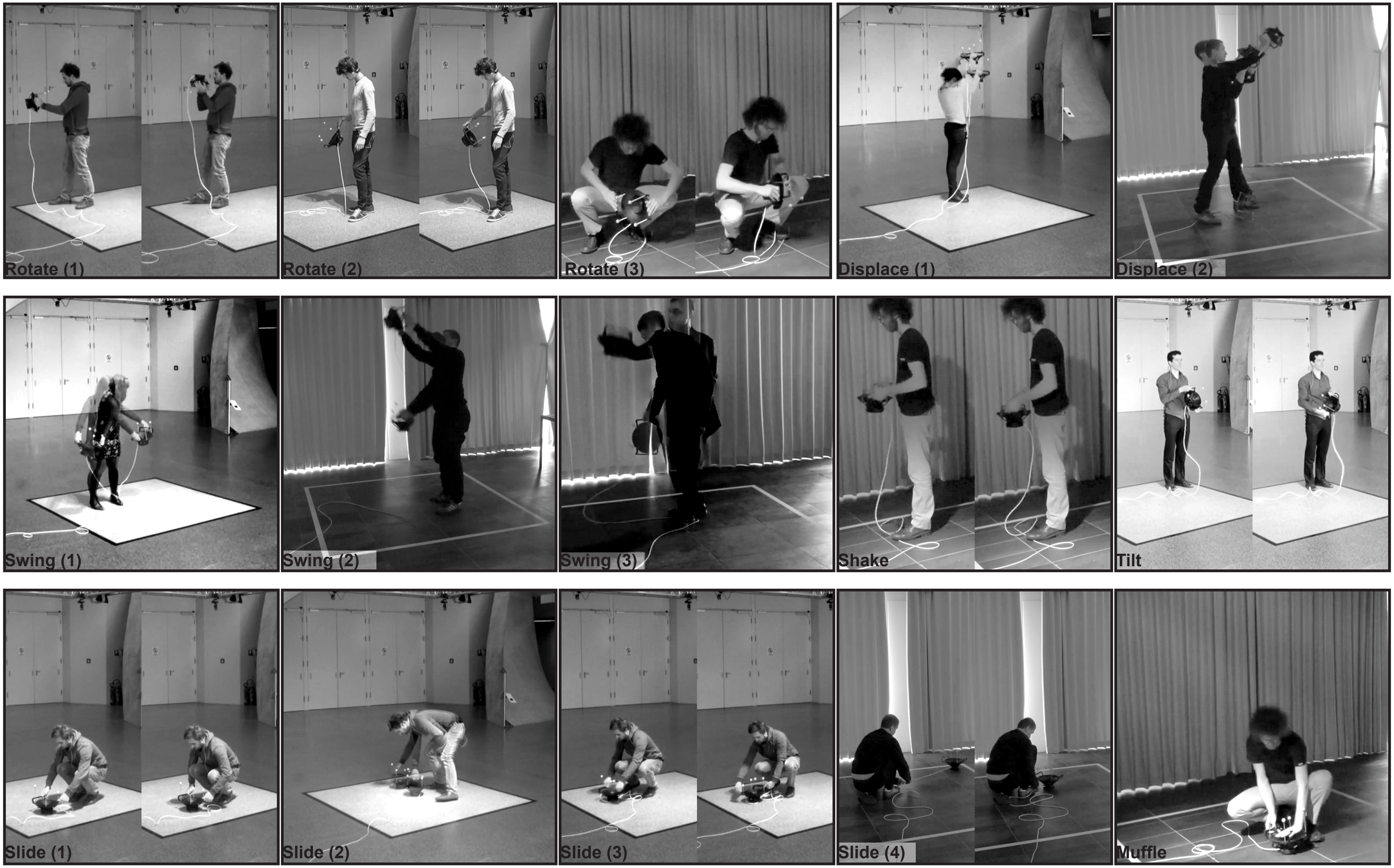
The different modes may well build on similar gesture combinations. When overlapping in form, the difference in the audible outcome is due to differences in the expressive quality of the gestures as this is shaped by the intention of the visitor. Laban effort analysis is used to describe these differences. The overlap is highest between crossing and tracing on the one hand and looping and manipulating on the other.

Crossing refers to instances in which participants cross the tracks laid out on the grid lines comprising the installation. This has the effect of hearing short segments of mixed sonic material. Perceptually, the effect is quite similar to quickly turning the radio tuner knob. Crossing functions as the primary searching and browsing mode and allows users to explore and return to areas of interest. Crossing involves primarily linear and to a lesser extent

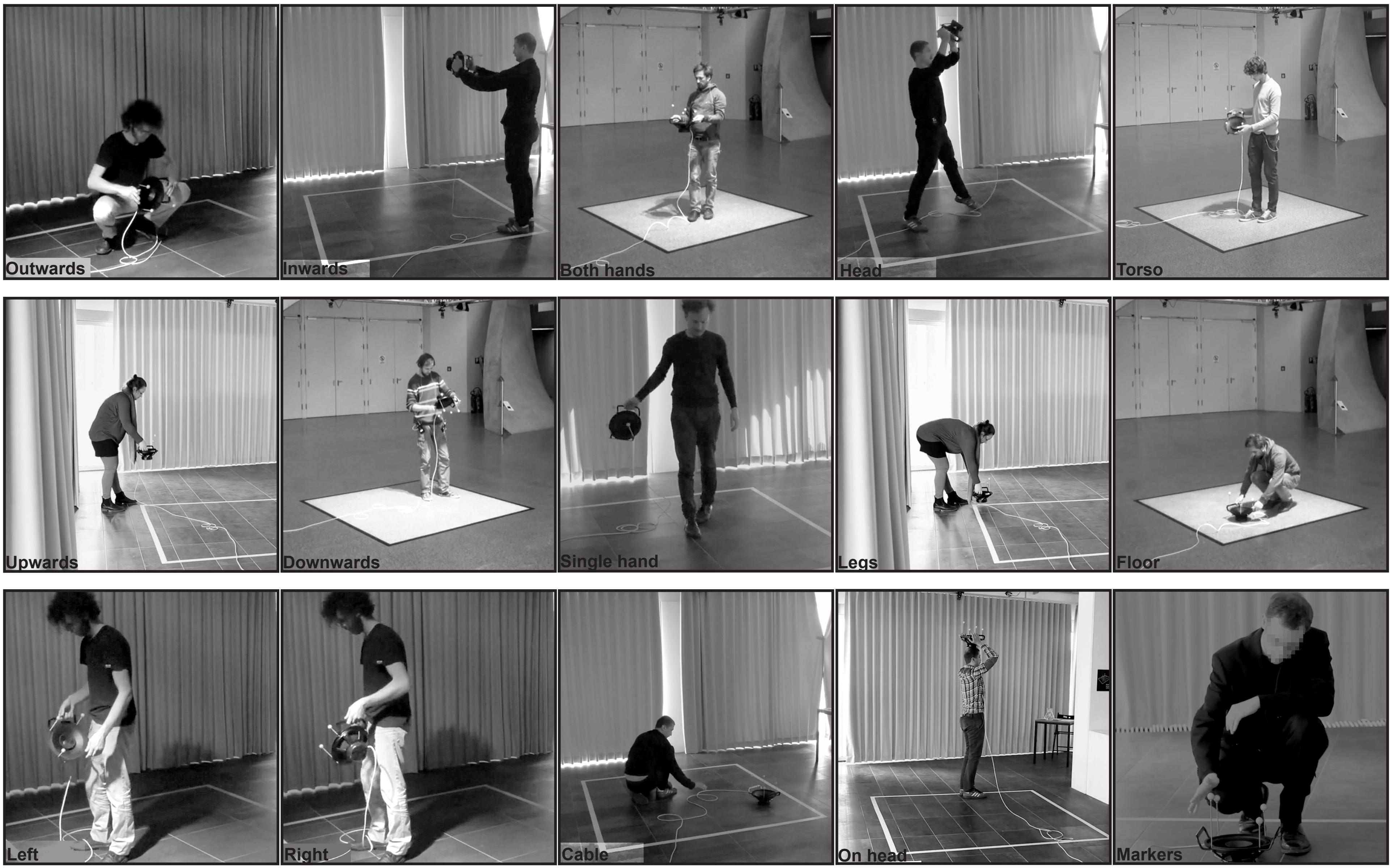
Body Gestures



Hand Gestures



Loudspeaker Orientation **Grip** **Hand Gesture Location**



circular or rotational loudspeaker movements of different amplitudes and speeds. Linear crossing is performed using **stand and displace** and **step and hold** gestures, followed by **step and displace**. In general, the space dimension is rather indirect both when moving the loudspeaker with the arms while standing but also when stepping. Visitors do not have a sharp spatial intention as they are searching and browsing in the absence of visual cues. There were a few exceptions when repositioning to a particular target or walking along the square marking the installation limits. Time and weight are neutral when crossing. When crossing is performed using **hold and turn** gestures a Wring quality emerges, indirect in space, sustained in time, and strong in weight. In a few cases, crossing was performed with an improvising intention, in which case movements remained indirect in space but are faster and more forceful, closer to the Slash movement quality. Some further notable combinations are stand or step and tilt (or swing), knee and displace (or slide), bend and displace (or hold), step and tilt (or rotate).

Tracing refers to instances in which users trace a track that has been laid upon a grid line. This has the perceptual effect of listening to a single voice. Depending on the amplitude of movement users may trace a phoneme, syllable, word, sentence, or part of a recitation. Tracing is used as means of exploring the detail of the sound tracks laid out in the installation. Some mastery of performing the installation is necessary before tracing gestures start to appear. Overall, to qualify as a tracing gesture the sound heard should be long enough to recognizable at least as a phoneme. Tracing is most commonly performed using **stand and displace** and **step and hold** gestures, however the frequency of step and displace gestures is lower compared to crossing. Overall, there is a tendency to keep the body stable while tracing. Other common combinations are stand and swing or tilt, knee and rotate, tilt or slide, bend and displace or hold, and hold and turn. Tracing involves primarily linear loudspeaker movements, however, for short sounds it may be done using circular or rotational movements. Space is now more direct as visitors attempt to trace a grid line accurately, time is sustained as they try to move carefully at a constant speed that makes text intelligible, and weight is light to allow for precise control. This is closest to the Glide movement quality.

Looping refers to instances in which users repeated the sound producing gesture continuously. This has the perceptual effect that a given entity (phoneme, syllable, word, or sentence) is heard repeatedly as a loop. Due to the design of the installation, this also had the (side-)effect that the sound written in the reverse direction may also be heard simultaneously depending on the speed with which the homing movement is performed. Looping allows to better focus on a track part but importantly looping also allows users to engage with the installation in a more creative way. Looping is typically performed using **displace gestures**, then by **tilting**, in a **standing or kneeing position**, and then by **holding from a standing position** or by **balancing back and forth**. The frequency of step and hold movements is lower than in tracing or crossing and balance and hold appears more often. As looping requires precision in repeating the gesture at a specific location space is direct especially when displacing or holding. When tilting, less attention is paid to space as movement is constrained by the gesture. Time is typically neutral, as is weight. Glide, Press, and Dab movement qualities are observed.

Modulating refers to instances in which users repeat their movements in a way that changes the sound color and creates a new sound that cannot be recognized as a phoneme. This is typically achieved through very fast or, less frequently, very slow looping. Modulating is an example of creative engagement with the installation as novel sounds are created. Most (about 75%) of Modulating excerpts are done using the **tilt gesture** from a static **standing**, **kneeing**, or less often **bending** body position using small amplitude (micro-movements) and high speed. Shaking from a standing position is also used as a modulating gesture, while rotating, swinging, displacing, and muffling follow. For Modulating, the dominating expressive quality involves direct space, quick time, and strong weight (Punch quality). In the less often slow movement cases, we have indirect space, sustained time, and strong weight (Press).

DISCUSSION

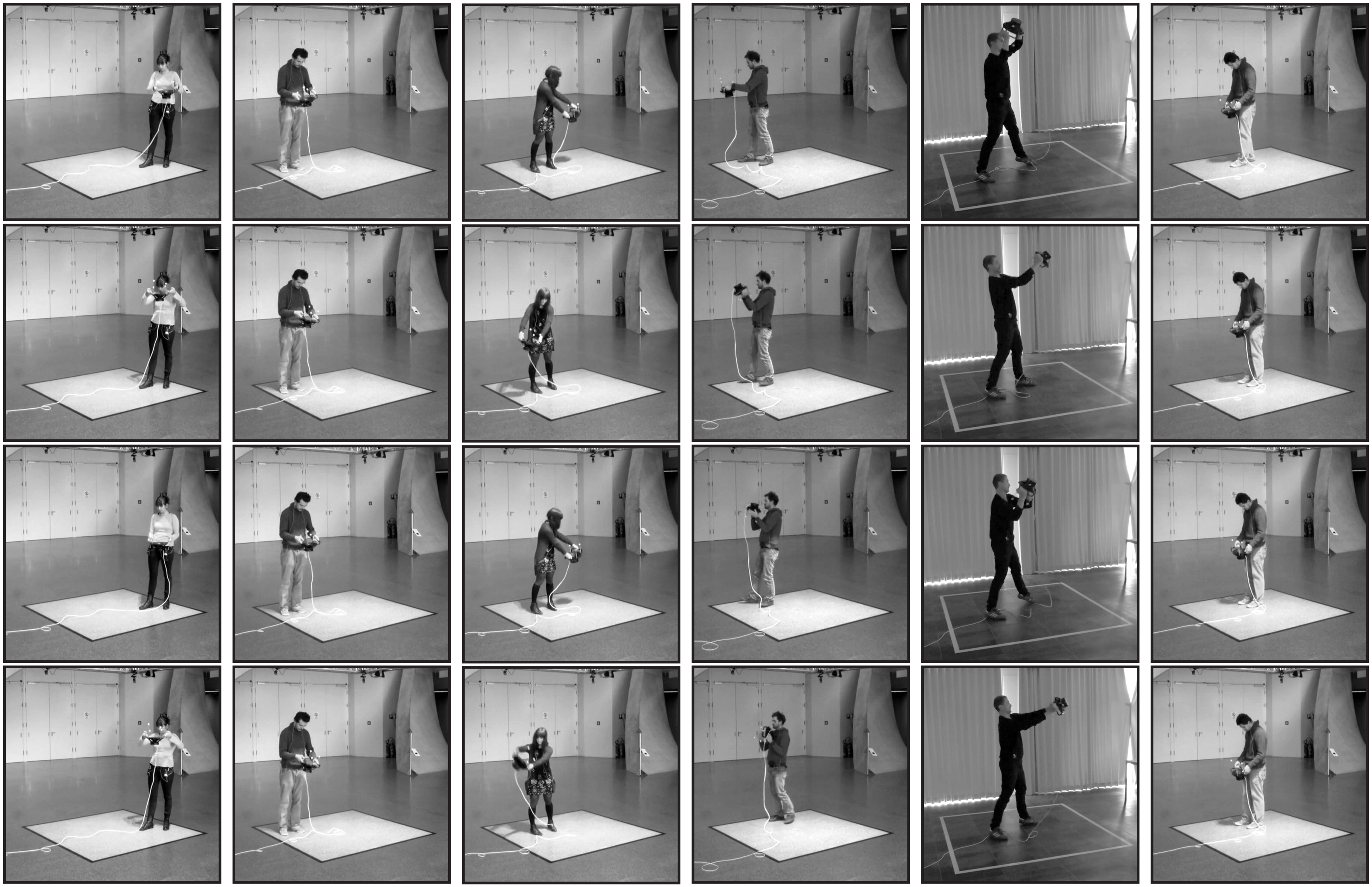
In search of ways to explore the design knowledge relating to movement-based interaction with sound embodied in the sound installation ‘Random Access Lattice’, we performed an inductive analysis of video recordings of the movements of evaluation participants. The installation lays out audio tracks as a lattice with axes parallel to the three cardinal axes. Visitors use a loudspeaker whose movement was tracked to interact with the sound laid out on the grid axes. The level of engagement displayed by participants during the exploration phase and the number of interactions we observed may suggest that this interaction paradigm was well received.

Form

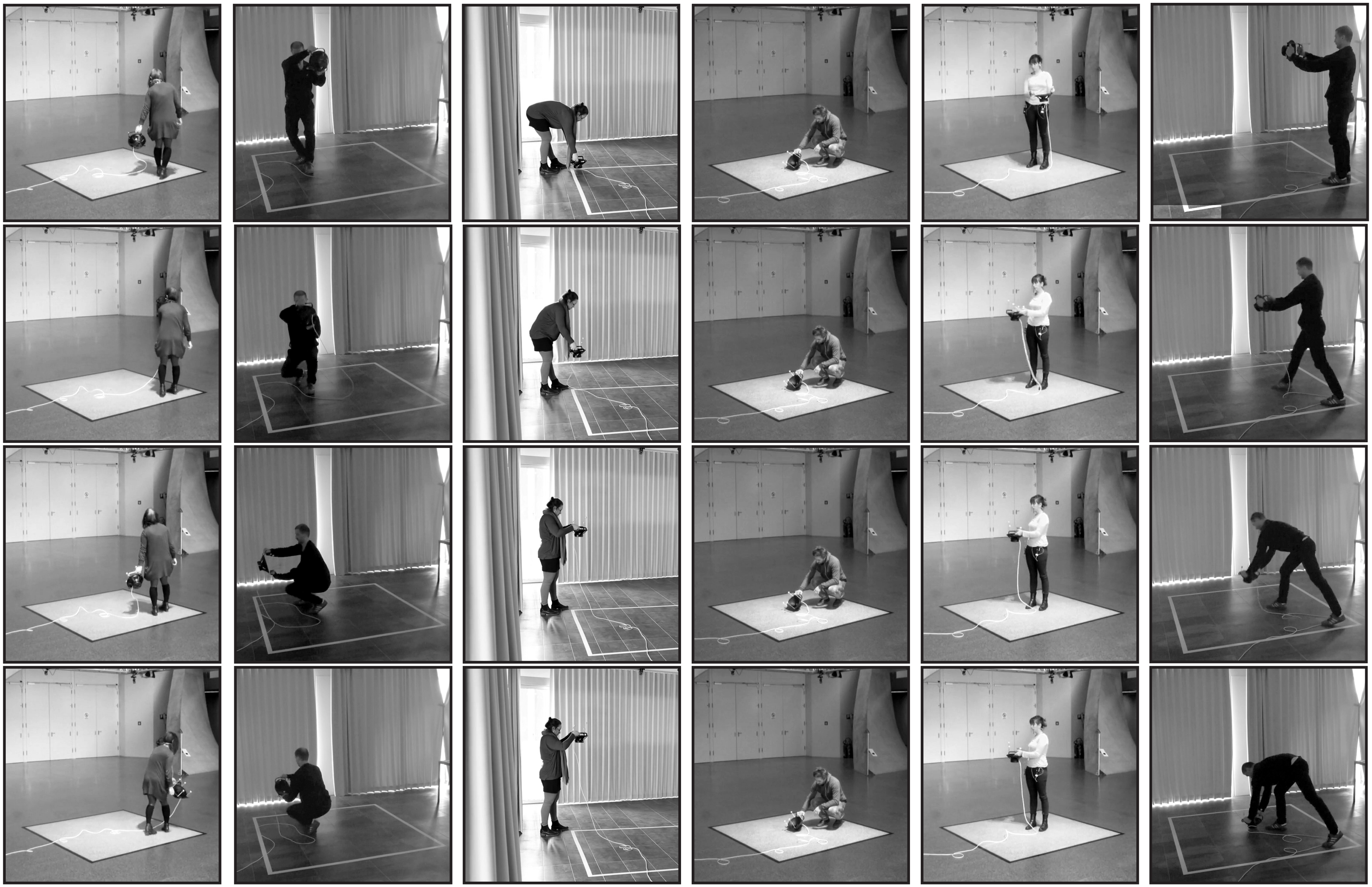
The gestures enacted by participants involved the whole body. We identified two defining aspects of gesture form. The first is the relationship between the body gesture, the arm gesture, and the hand gesture. The second is the type of movement. Two movement types were observed, linear and circular, depending on whether the loudspeaker moved along a linear or a curved path in 3D space. Both movement types can be performed from both static and dynamic body and arm positions. Most commonly, linear movements appear in relation to displace and hold arm gestures. Circular movements are related to the tilt or rotate arm gestures in dynamic body positions and hold gesture while turning the body.

The dynamics of body and arm gestures are quite opposite. Static body gestures are most often combined with dynamic arm gestures and vice-versa. For example, displace, tilt, or swing gestures are most often performed from a standing position while the hold gesture is most common while stepping, balancing, or turning the body. This

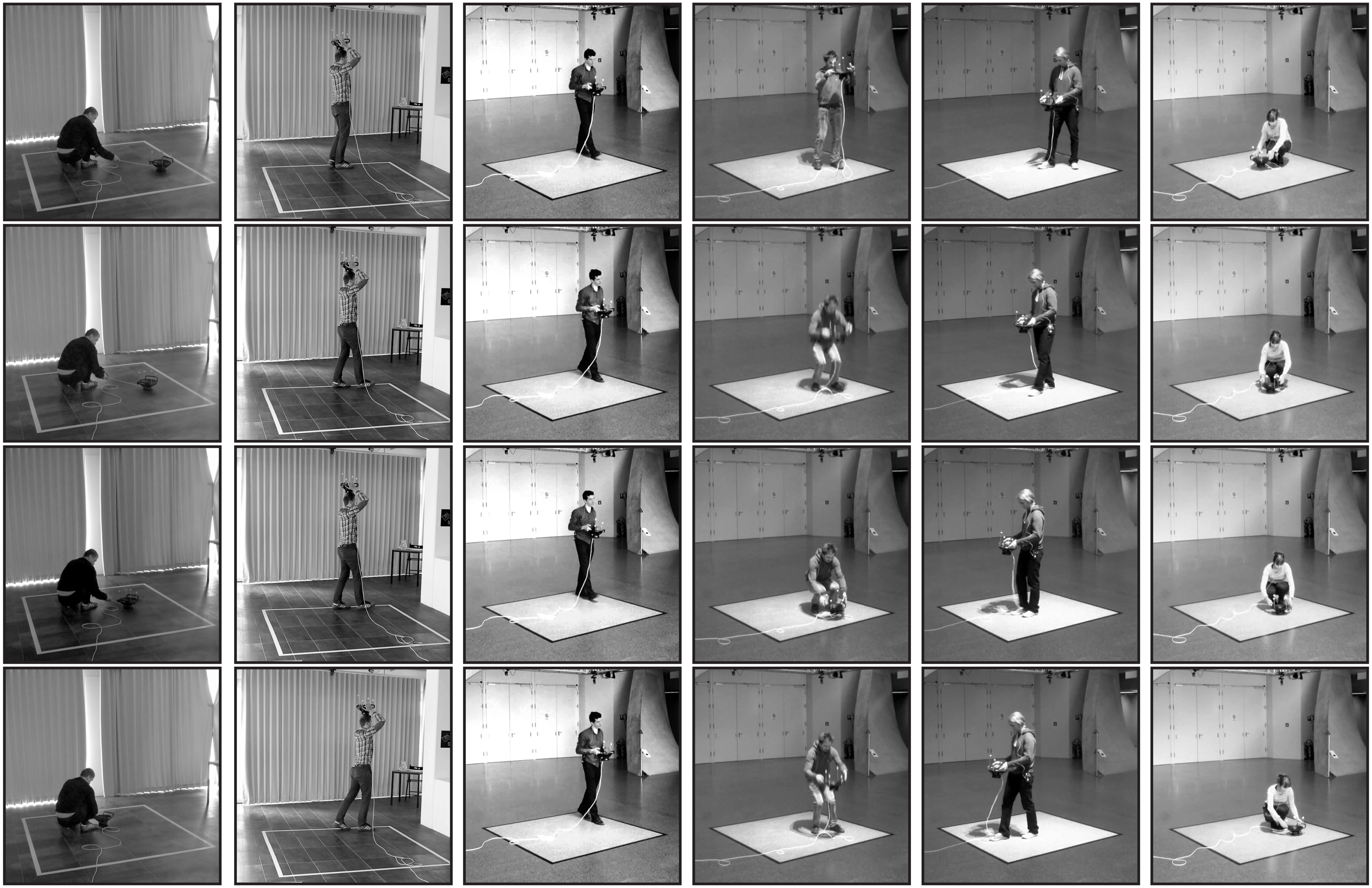
Stand and Displace | **Stand and Tilt** | **Stand and Swing** | **Stand and Rotate** | **Step and Displace** | **Step and Tilt**



Step and Swing | **Knee and Hold** | **Bend and Displace** | **Knee and Tilt** | **Turn and Hold** | **Bend and Hold**



Knee and Slide (Cable) | **Step and Hold (Head)** | **Balance and Hold** | **Knee and Displace** | **Step and Hold** | **Knee and Slide**



dipole relates to interacting with sonic material in the personal space versus seeking or accessing sonic material in remote locations.

Expression

Expressive movement qualities allow us to differentiate between gestures of similar form. As mentioned earlier, similar gestures were used when crossing and tracing on the one hand and when looping and manipulating on the other. Laban Effort-Shape qualities were instrumental in describing the differences in gesture performance that led to different interactions as a result of gestures of similar form. Despite having similar form, crossing gestures were more diffuse in space and time than tracing gestures. Time was neutral for looping but the likelihood of slow and fast time increased for modulating. Overall large movement amplitudes led to less spatial precision. Crossing involves the largest movement amplitude, and was most diffuse, followed by tracing, and then looping, and modulating. Diffuseness in space can also be related to the weighting of the sonic output for the motor performance of the gesture. While this is low for crossing and gestures lean towards an open-loop behavior, it is high for tracing, looping, and modulating, as these are based on exact movements and resemble closed-loop behavior. Overall, the weight dimension was heavy as a result of holding the loudspeaker. Flow was also typically bound as visitors tried to control the loudspeaker. There were a few instances with free flow when crossing around the installation.

Sonic Affordances

The four interaction modes we identified can be understood as the primary affordances offered by the installation. Crossing is about exploring, locating, and returning to sounds. Linear and curved loudspeaker movement trajectories are used. The lack of visual references makes it unavoidable that crossing is ephemeral and inexact. Only small parts of sound (previews) are revealed while crossing. Tracing is about listening. Meaning in auditory experiences unfolds serially in time and tracing is there to support such a listening experience. It is about precision, timing, and respecting the temporal order of the sonic material as this has been engraved into space. Linear movement trajectories along the lattice lines dominate. Often, movements are repeated in order to increase precision. This leads to looping, a serendipitous discovery, which often emerges in the attempt to get tracing right. But looping allows focusing on the looped part and brings out emergent sonic qualities and effects depending on the gesture dynamics. Curved movement trajectories appear more often and the frequency of tilting increases as a convenient way to perform repetitive movements accurately. This behavior is accentuated when modulating in which case curved movement trajectories are very often.

Importantly, what we find is that participants use variations in both form and expression to enact these affordances. Form differences are more pronounced when contrasting crossing and tracing to looping and modulating. There, there is a tendency to move from linear (displace, hold) to curved movement trajectories (tilt, rotate); an interesting reference to the increased periodicity of the sonic outcome. A telling exception is looping using the balancing movement. Here, participants engage with a linear activity like stepping in a repetitive periodic way (balancing) in order to enact looping. Differences in expression are most pronounced when contrasting crossing to tracing and looping to modulating. By manipulating Laban Effort qualities, participants can accomplish the finer tasks of listening to the sonic material or modulating the sound timbre.

One aspect that needs to be discussed is the affordances of the loudspeaker design and the presence of the cable. In the RAL realization we tested, a sizable loudspeaker driver was used with two handles attached to it. Visitors manipulated the loudspeaker while grasping it by the handles. Furthermore, there was a cable which connected the loudspeaker to the computer running the software for the installation. This setup obviously shaped the affordances perceived by visitors as handles were used extensively in nearly all gestures we observed. Handles were used creatively in gestures involving both hands but also a singe hand and were vital to allowing precise control of the loudspeaker. Concerning the cable, some participants commented that it was an annoyance, while others were not bothered by its presence, or even perceived it as an affordance as when sliding the loudspeaker on the floor. In the meanwhile, the artist has developed a variation of the installation in which a wireless loudspeaker is used and in future work it would be interesting to observe how adds to the observations reported here.

Quite clearly, both movement form and expression need to be considered when designing gesture interaction with sound. The importance of movement quality is fundamental in musical expression. However, this aspect has not been investigated explicitly when designing sonic interactions, as designers focus mostly on the gesture form and less on expressive qualities. In line with the existing discourse in designing for movement [1], the application of relevant design methodologies is necessary when designing sonic interactions in order to experience, design, specify and associate movement qualities with sonic outcomes in sufficient detail to create fine interactions for searching, browsing, listening, looping, and modulating sound.

ACKNOWLEDGMENTS

The author would like to acknowledge the invaluable support of David Pirrò, Marian Weger, and Nick Acorne for realizing the evaluation and obtaining the video recordings. The author is grateful to Gerhard Eckel for sharing RAL for evaluation, material describing the installation, and his views during helpful discussions. This work was supported by the Klangräume project (Zukunftsfonds Steiermark PN:6067).

REFERENCES

[1] Sarah Fdili Alaoui, Baptiste Caramiaux, Marcos Serrano, and Frédéric Bevilacqua. 2012. Movement qualities as interaction modality. In Proceedings of the Designing Interactive Systems Conference. 761–769.

[2] Alessandro Altavilla, Baptiste Caramiaux, Atau Tanaka, et al. 2013. Towards gestural sonic affordances. (2013).

[3] Stephen Barrass. 2013. ZiZi: The Affectionate Couch and the Interactive Affect Design Diagram. Sonic Interaction Design (2013), 235.

[4] Ilias Bergström and Martin Jonsson. 2016. Sarka: sonification and somaesthetic appreciation design. In Proceedings of the 3rd International Symposium on Movement and Computing. ACM, 1.

[5] Ilias Bergstrom, Sofia Seinfeld, Jorge Arroyo-Palacios, Mel Slater, and Maria V Sanchez-Vives. 2014. Using music as a signal for biofeedback. International Journal of Psychophysiology 93, 1 (2014), 140–149.

[6] Ulysses Bernardet, Sarah Fdili Alaoui, Karen Studd, Karen Bradley, Philippe Pasquier, and Thecla Schiphorst. 2019. Assessing the reliability of the Laban Movement Analysis system. PloS one 14, 6 (2019), e0218179.

[7] Zafer Bilda. Evaluating audience experience with interactive art. Engage, 2006.

[8] Stephen Brewster, Joanna Lumsden, Marek Bell, Malcolm Hall, and Stuart Tasker. 2003. Multimodal'eyes-free'interaction techniques for wearable devices. In Proceedings of the SIGCHI conference on Human factors in computing systems. 473–480.

[9] Vincent Buil, Gerard Hollemans, and Sander van deWijdeven. 2005. Headphones with touch control. In Proceedings of the 7th international conference on Human computer interaction with mobile devices & services. ACM, 377–378.

[10] Linda Candy. 2014. Interactive Experience in the Digital Age. Springer, Chapter Evaluation and Experience in Art, 25–48.

[11] Baptiste Caramiaux, Frédéric Bevilacqua, Tommaso Bianco, Norbert Schnell, Olivier Houix, and Patrick Susini. 2014. The role of sound source perception in gestural sound description. ACM Transactions on Applied Perception (TAP) 11, 1 (2014), 1.

[12] Stefano Delle Monache, Davide Rocchesso, Frédéric Bevilacqua, Guillaume Lemaitre, Stefano Baldan, and Andrea Cera. 2018. Embodied sound design. International Journal of Human-Computer Studies 118 (2018), 47–59.

[13] Lina Dib, Daniela Petrelli, and Steve Whittaker. 2010. Sonic souvenirs: exploring the paradoxes of recorded sound for family remembering. In Proceedings of the 2010 ACM conference on Computer supported cooperative work. 391–400.

[14] Paul Dourish. 2004. Where the action is: the foundations of embodied interaction. MIT press.

[15] Tijs Duel, David M Frohlich, Christian Kroos, Yong Xu, Philip JB Jackson, and Mark D Plumbley. 2018. Supporting audiography: Design of a system for sentimental sound recording, classification and playback. In International Conference on Human-Computer Interaction. Springer, 24–31.

[16] Gerhard Eckel. 2013. Improvisation as epistemic practice. Malmö Academy of Music, Lund University, 42–47.

[17] Sarah Fdili Alaoui, Jules Françoise, Thecla Schiphorst, Karen Studd, and Frederic Bevilacqua. 2017. Seeing, sensing and recognizing Laban movement qualities. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 4009–4020.

[18] Sarah Fdili Alaoui, Thecla Schiphorst, Shannon Cuykendall, Kristin Carlson, Karen Studd, and Karen Bradley. 2015. Strategies for embodied design: The value and challenges of observing movement. In Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition. ACM, 121–130.

[19] Jules Françoise, Yves Candau, Sarah Fdili Alaoui, and Thecla Schiphorst. 2017. Designing for kinesthetic awareness: Revealing user experiences through second-person inquiry. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 5171–5183.

[20] Jules Françoise, Sarah Fdili Alaoui, Thecla Schiphorst, and Frédéric Bevilacqua. 2014. Vocalizing dance movement for interactive sonification of laban effort factors. In Proceedings of the 2014 conference on Designing interactive systems. 1079–1082.

[21] Karmen Franinovic and Stefania Serafin. 2013. Sonic interaction design. MIT Press.

[22] Dennis Gabor. 1946. Theory of communication. Part 1: The analysis of information. Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering 93, 26 (1946), 429–441.

[23] William Gaver. 2012. What should we expect from research through design?. In Proceedings of the SIGCHI conference on human factors in computing systems. ACM, 937–946.

[24] Rolf Inge Godøy, Egil Haga, and Alexander Refsum Jensenius. 2005. Playing “air instruments”: mimicry of sound-producing gestures by novices and experts. In International Gesture Workshop. Springer, 256–267.

[25] T. Hermann, A. Hunt, and John G. Neuhoff. 2011. The sonification handbook. Logos Verlag Berlin, GE.

[26] Mads Hobye and Jonas Löwgren. 2011. Touching a stranger: Designing for engaging experience in embodied interaction. International Journal of Design 5, 3 (2011), 31–48.

[27] Kristina Höök, Baptiste Caramiaux, Cumhur Erkut, Jodi Forlizzi, Nassrin Hajinejad, Michael Haller, Caroline Hummels, Katherine Isbister, Martin Jonsson, George Khut, et al. 2018. Embracing first-person perspectives in soma-based design. In Informatics, Vol. 5. Multidisciplinary Digital Publishing Institute, 8.

[28] Kristina Höök and Jonas Löwgren. 2012. Strong concepts: Intermediate-level knowledge in interaction design research. ACM Transactions on Computer-Human Interaction (TOCHI) 19, 3 (2012), 23.

[29] Kristina Höök, Phoebe Sengers, and Gerd Andersson. 2003. Sense and sensibility: evaluation and interactive art. In Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 241–248.

[30] Eva Hornecker and Jacob Buur. 2006. Getting a grip on tangible interaction: a framework on physical space and social interaction. In Proceedings of the SIGCHI conference on Human Factors in computing systems. 437–446.

[31] Sergi Jordà, Günter Geiger, Marcos Alonso, and Martin Kaltenbrunner. 2007. The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. In Proceedings of the 1st international conference on Tangible and embedded interaction. 139–146.

[32] Minoru Kobayashi and Chris Schmandt. 1997. Dynamic Soundscape: mapping time to space for audio browsing. In SIGCHI. ACM Press, Atlanta, Georgia, United States, 194 – 201.

[33] Wonjun Lee, Youn-kyung Lim, and Richard Shusterman. 2014. Practicing somaesthetics: exploring its impact on interactive product design ideation. In Proceedings of the 2014 conference on Designing interactive systems. ACM, 1055–1064.

[34] Guillaume Lemaitre, Olivier Houix, Yon Visell, Karmen Franinović, Nicolas Misdariis, and Patrick Susini. 2009. Toward the design and evaluation of continuous sound in tangible interfaces: The Spinotron. International Journal of Human-Computer Studies 67, 11 (2009), 976–993.

[35] Yueh-Tuan Li and Wen-Shu Lai. 2013. Voice, object and listening in the sound installations of Laurie Anderson. International Journal of Performance Arts and Digital Media 9, 2 (2013), 347–361.

[36] Lian Loke, Astrid T Larssen, and Toni Robertson. 2005. Labanotation for design of movement-based interaction. In Proceedings of the second Australasian conference on Interactive entertainment. Creativity & Cognition Studios Press, 113–120.

[37] Lian Loke and Toni Robertson. 2009. Design representations of moving bodies for interactive, motion-sensing spaces. International Journal of Human-Computer Studies 67, 4 (2009), 394–410.

[38] Lian Loke and Toni Robertson. 2013. Moving and making strange: An embodied approach to movement-based interaction design. ACM Transactions on Computer-Human Interaction (TOCHI) 20, 1 (2013), 7.

[39] Jonas Löwgren. 2013. Annotated portfolios and other forms of intermediate-level knowledge. interactions 20, 1 (2013), 30–34.

[40] G. Marentakis and S. Brewster. 2006. Effects of feedback, mobility and index of difficulty on deictic spatial audio target acquisition in the horizontal plane. In Proceedings of the SIGCHI conference on Human Factors in computing systems. ACM, 359–368.

[41] G Marentakis and Stephen McAdams. 2013. Perceptual impact of gesture control of spatialization. ACM Transactions on Applied Perception (TAP) 10, 4 (2013), 1–21.

[42] G. Marentakis, D. Pirrò, and M. Weger. 2017. Creative Evaluation. In Proceedings of the 2017 Conference on Designing Interactive Systems. ACM, 853–864. https://doi.org/10.1145/3064663.3064710

[43] Andrew P McPherson, Alan Chamberlain, Adrian Hazzard, Sean McGrath, and Steve Benford. 2016. Designing for exploratory play with a hackable digital musical instrument. In Proceedings of the 2016 ACM Conference on Designing Interactive Systems. 1233–1245.

[44] Helena M Mentis and Carolina Johansson. 2013. Seeing movement qualities. In Proceedings of the sigchi conference on human factors in computing systems. 3375–3384.

[45] Christian Metzger, Matt Anderson, and Thad Starner. 2004. Freedigiter: A contact-free device for gesture control. In Wearable Computers, 2004. ISWC 2004. Eighth International Symposium on, Vol. 1. IEEE, 18–21.

[46] Jos Mulder. 2010. The loudspeaker as musical instrument. In Proceedings of the 2010 Conference on New Interfaces for Musical Expression. 13–18.

[47] Gerard Oleksik, David Frohlich, Lorna M Brown, and Abigail Sellen. 2008. Sonic interventions: understanding and extending the domestic soundscape. In Proceedings of the SIGCHI conference on Human Factors in computing systems. 1419–1428.

[48] Nam June Paik. 2002. Random Access. Leonardo 35, 5 (2002), 489–489.

[49] Marco Pasch, Nadia Bianchi-Berthouze, Betsy van Dijk, and Anton Nijholt. 2009. Immersion in movement-based interaction. In International Conference on Intelligent Technologies for Interactive Entertainment. Springer, 169–180.

[50] Daniela Petrelli, Nicolas Villar, Vaiva Kalnikaite, Lina Dib, and Steve Whittaker. 2010. FM radio: family interplay with sonic mementos. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2371–2380.

[51] Antti Pirhonen, Stephen Brewster, and Christopher Holguin. 2002. Gestural and audio metaphors as a means of control for mobile devices. In Proceedings of the SIGCHI conference on Human factors in computing systems. 291–298.

[52] M. Rumori and G. Marentakis. 2017. Paris Flâneur - An artistic approach to binaural technology. Presence: Teleoperators & Virtual Environments 26, 2 (2017), 111–137. https://doi.org/10.1162/PRES_a_00289

[53] N. Sawhney and C. Schmandt. 2000. Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. ACM transactions on Computer-Human interaction (TOCHI) 7, 3 (2000), 353–383.

[54] Bert Schiettecatte and Jean Vanderdonckt. 2008. AudioCubes: a distributed cube tangible interface based on interaction range for sound design. In Proceedings of the 2nd international conference on Tangible and embedded interaction. 3–10.

[55] Thecla Schiphorst. 2011. Self-evidence: applying somatic connoisseurship to experience design. In CHI'11 extended abstracts on human factors in computing systems. ACM, 145–160.

[56] Lisa Sifelman, Bary Arons, and Chris Schmandt. 2001. The Audio Notebook. Paper and Pen Interaction with Structured Speech. In SIGCHI 01, Vol. 3(1). ACM, Seattle, WA, USA, 182–189.

[57] Vidmina Stasiulyte. 2018. The Aesthetics of the Invisible: Foundations for Sonic Fashion. Ph.D. Dissertation. Högskolan i Borås.

[58] Ana Tajadura-Jiménez, Aleksander Väljamäe, Frédéric Bevilacqua, and Nadia Bianchi-Berthouze. 2018. Principles for Designing Body-Centered Auditory Feedback. The Wiley Handbook of Human Computer Interaction Volume (2018), 371.

[59] Atau Tanaka, Alessandro Altavilla, and Neal Spowage. 2012. Gestural Musical Affordances. Proceedings of the 9th International Conference on Sound and Music Computing, Copenhagen, Denmark, 2012.

[60] Atau Tanaka and Marco Donnarumma. 2019. The body as musical instrument. The Oxford Handbook of Music and the Body (2019), 79.

[61] Theophanis Tsandilas. 2018. Fallacies of agreement: A critical review of consensus assessment methods for gesture elicitation. ACM Transactions on Computer-Human Interaction (TOCHI) 25, 3 (2018), 1–49.

[62] Yolanda Vazquez-Alvarez, Ian Oakley, and Stephen A Brewster. 2012. Auditory display design for exploration in mobile audio-augmented reality. Personal and Ubiquitous computing 16, 8 (2012), 987–999.

[63] Marcelo M Wanderley and Philippe Depalle. 2004. Gestural control of sound synthesis. Proc. IEEE 92, 4 (2004), 632–644.

[64] John Williamson, Roderick Murray-Smith, and Stephen Hughes. 2007. Shoogle: excitatory multimodal interaction on mobile devices. In Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 121–124.

[65] Todd Winkler. 1995. Making motion musical: Gesture mapping strategies for interactive computer music. In ICMC. 26–30.