AR Sound Sandbox: A Playful Interface for Musical and Artistic Expression

Bastian Dewitz[1], Roman Wiche[1], Chris Geiger[1],
Frank Steinicke[2], and Jochen Feitsch[1](✉)

[1] Duesseldorf University of Applied Sciences, Duesseldorf, Germany
{bastian.dewitz,geiger,jochen.feitsch}@hs-duesseldorf.de,
roman.w92@web.de
[2] University of Hamburg, Hamburg, Germany
steinicke@informatik.uni-hamburg.de

Abstract. In this paper we introduce a novel interface combining spatial and continuous tangible interaction for creating and manipulating audio-visual effects. Our goal is to provide a ready-to-use, "hands-on" interface that does not need protracted explanation to the user, yet provides possibilities for expression. Therefore, our interface exploits the three-dimensional topology of physical sand, which is distributed over a tabletop surface. We discovered, that users of the system were engaged by the natural interaction and playful manner of the installation, as it resembles the play in a sandbox. We demonstrate an artistic setup that produces ambient soundscapes using a Lattice Boltzmann based particle simulation running through a deformable landscape. Visual feedback is front-projected onto the sand as well as the user's hand. The user can explore and change the landscape by using his or her hands and use spatial gestures via on-body projection to control AR content and further settings. The focus of this work lies on the simultaneous interaction with sand and the user's own body, and it's contribution to audio-visual installations.

Our prototype system was tested with potential users in a small informal study and was overall well received. Users had fun exploring the different forms of interaction techniques to control the particle simulation and soundscape, and were amazed by the possibilities of on-body interaction. In future we plan to further evaluate our system in a formal study and compare interaction and user experience to similar interfaces. The system was successfully deployed as an indoor room installation, reducing it's components to a minimum. Further deployments are planned.

Keywords: Musical human-computer interaction
Haptic and force feedback devices
Interactive sound art and installations
Novel controllers and interfaces for musical expression
Continuous tangible user interfaces · On-body interaction
Spatial user interface · Projection mapping

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2018
Y. Chisik et al. (Eds.): INTETAIN 2017, LNICST 215, pp. 59–76, 2018.
https://doi.org/10.1007/978-3-319-73062-2_5

1 Introduction

Despite the advances in technologies for interacting with computer systems over the last years, operating computers is still mostly based on mouse, keyboard or recently touch screens. This form of interaction differs from the natural interaction with the real environment. Therefore, researchers have been investigating for the last decades how new forms of human computer interaction can become more similar to the interaction with the real world.

One approach to tackle this problem is investigated in the research field of tangible user interfaces (TUI). The basic idea behind TUIs is to use natural interaction with real objects in the real environment to control digital content. The mapping between physical objects and digital content is done in different ways, for example, by combining them modularly or by arranging them spatially. Objects that map physical manipulation onto digital data are called tangibles.

Especially the research area of new musical interfaces frequently yields new prototype systems for musical expression that adopt HCI-concepts for natural and tangible interaction. Some of these prototypes can be rather successful, for example reacTable, which was developed since 2003 by Jordà [6] and can now even be used as a new instrument in live performances. Interaction techniques and metaphors vary strongly between prototypes so a vast quantity of different systems exists. Qualitative user studies show promising results for new forms of collaborative interaction, user experience and entertainment [1,6,7]. However, a rarely used approach for musical interfaces are continuous tangible user interfaces, a concept of TUIs developed in 2002 by MIT media laboratory [9].

In our approach for a new way of musical expression we transfer the sculpting of a landscape to the sculpting of a soundscape. We combine tangible interaction, interactive augmented reality (AR) content and on-body interaction in our system to create a novel way of interacting with music and sound. The user can interact with the system by using only his or her hands and does not need any additional input devices. Therefore, the system is ready-to-use: once set up, users can interact with an installation without further requirements. We developed a prototype to evaluate this concept for interaction in a user study to identify possible applications and potentials. While the basic interaction of sculpting the sand is self-explanatory, the advanced interaction techniques require a short explanation.

The remainder of this paper structures as follows: First we give an overview of continuous tangible user interfaces for musical expression and on-body interaction prototypes that have been developed in the last years. Then we present design concepts and implementation details of our own approach. We proceed by presenting observations made during a qualitative user study conducted with our system. We conclude with a discussion of these observation and plans on further development and evaluation of our system.

2 Related Work

We combine two new approaches for spatial human-computer interaction in our system: continuous tangible user interfaces and on-body interaction. In the following we give an outline of continuous tangible user interfaces for musical expression and on-body interaction prototype systems.

2.1 Interfaces for Musical Expression

There are plenty examples of TUI based prototype systems for musical expression: A common approach is to utilize tangible artifacts to control certain aspects of music, for example, fading volume by turning a tangible. The spatial arrangement of these artifacts is used for interaction with music and influences the behavior of spatially close artifacts. By this, the tangibles become a metaphor of the music itself, controlling main parameters and influencing each other. Research results show that these kind of interfaces are generally well received and both easy and fun to use [1,6], in some aspects facilitate interaction, for example, concerning multi-user-interaction, and allow for a playful exploration of the interface [6]. However, there are just very few prototype systems that explore the application of continuous tangible user interfaces as interaction concept:

Granulatsynthese was an early attempt in 2008 to create a music interface using a continuous deformable surface [1]. It is highly focused on the individual user experience and not determined to enable the user to create sounds or music in a performance way. Beckhaus et al. used a granular half-transparent substance as a haptic input medium. The user would dig holes that were detected by an infrared light system. Depending on the operating preset and the created shapes the user would hear synthetic sounds and see matching computer generated sound-waves, which are projected onto to granules from the backside, resulting in a mesmerizing atmosphere [1]. Granulatsynthese was tested multiple times on different occasions and with different test subjects. Beckhaus et al. describe a playful interaction with the system and exploratory behavior. They especially point out the importance of the tangible medium as it significantly influences the haptic experience.

In 2014 the students project Sand Noise Device [11] features a sandbox as sound generating game-like device. Users can place tangible glowing pucks on a sand surface which is captured by a Kinect sensor. AR content is projected onto the sand from top. In fixed time intervals sound bursts are emitted from these pucks and a central emitter along with projected particles. These particles move through the sand landscape to the lowest point in their vicinity. The velocity of each particle is then mapped to a sound generation in MaxMSP. The system received some media attention and was overall well received but unfortunately neither further evaluated nor scientifically published.

A recent approach is soundFORMS [2], a synthesizer and sequencer created by MIT media Lab in 2016. It allows users to shape waveforms and drum-loops with a predefined set of gestures, for example, a chopping down movement with one hands to create a sawtooth waveform, which is then rendered as audio. The system uses a 24×24 matrix of motorized pins that change the appearance of the surface according to the created sound-waves.

2.2 On-Body Interaction Prototypes

The research field of on-body projection and interaction is a very young research area with a small number of prototypes that primarily demonstrate the feasibility of this new concept for interacting with digital content. The basic idea behind on-body interaction is to use the human body as input or output device by projecting computer generated interactive content onto specific body parts. Usually such a system consist of three components [5]: A tracking device that allows to determine the 3D-position of body parts in the real world, a projector that projects content onto these points and a computer that allows interaction with the projected content. In most cases, content is projected onto the users hands or arms as these body parts can be easily placed in field of view and provide our primary tool for interacting with the real world [5].

In the past few years, driven by the advancing technological possibilities, stationary [4,12] and wearable [3,5,8] prototypical systems have been successfully developed. Many systems rely on depth [3,4,12] or RGB [8] camera data and use computer vision algorithms to detect relevant body parts or movement. Depending on the usage scenario, different forms of interaction have been tested, in many cases some form of gestural input [4,8,12] or touch event based interaction with virtual AR buttons and sliders [3,5]. In some cases, the surrounding environment is used as an additional projection area [8,12]. Overall, despite the feasibility of the concept of on-body interaction, no best practices for technologies or interaction have been established yet.

3 Projection-Based AR Sound Sandbox

As there is no direct analogy between sculpting a sand surface and creating a soundscape, it is possible to freely design different approaches for interacting with sound. The concept in this approach was developed considering following criteria:

- Easy to understand reaction of the system to input of the user, allowing a deliberate manipulation of produced sound. The user controls the volume levels of different soundscape tracks.
- Correspondence between visual effects and audio.
- Combine AR content, on-body interaction and deformation of the sand surface to control the system.
- Create a system that is fun to interact with.

3.1 Setup

Our sandbox uses a setup similar to Piper and Ishiis Illuminating clay [9] respectively Sand Noise Device [11]. We use a 0.6 m × 0.4 m plexiglass box that is filled

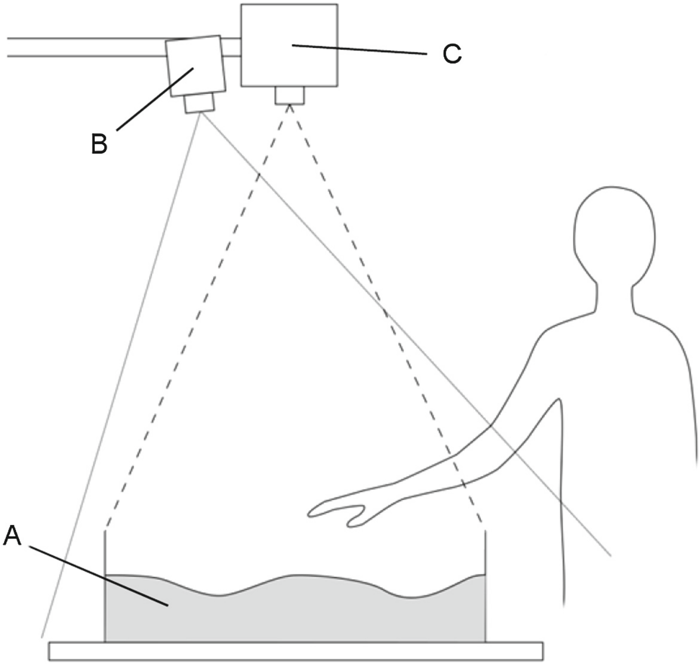

Fig. 1. Prototype setup. A: deformable sand surface, B: projector, C: depth sensor.

with 0.1 m kinetic sand. We chose kinetic sand as haptic medium because of its interesting haptic features and its easy formability into firm structures. To provide preferably genuine projection features, we used a special white-colored sand. We placed a Kinect 2 sensor and a small projector (Optoma ML750e) 1.2 m above the sand surface, both facing downwards (see Fig. 1). In this setup the Kinect 2 sensor provides a 184 × 120 pixels wide depth image of the sandbox and user hands positioned above the surface. The system we are using is a Windows 8.1 PC equipped with an Intel i7-3930k CPU, 8 GB Ram and a Geforce GTX 770. Our development environment is the game engine Unity 5.

3.2 Projection

To project AR content onto the deformable sand surface, we determine the transformation matrix between Kinect sensor space and projector image space. This allows rendering a background image onto the sand to colorize the surface in different color schemes. We use a dynamic grid that deforms accordingly to the sand's topography to correctly project a colored background image. First we define the outline of the sandbox in the Kinect depth image with the four vertices of a quad. This quad is then divided into a grid of 32 × 32 cells. Each vertex of these cells has a distinct 3D-position in Kinect space which can be transformed into 2D projector space coordinates to distort a textured mesh to match the sands topography, resulting in a approximated projection onto the surface (see Fig. 2). Using the same transformation matrix, additional AR-objects like 3D models or 2D images can be rendered onto specific points of the sands surface or the users hand.

Fig. 2. Distorted grid projected onto the sand surface

We aimed for a harmonized presentation of sound and visual effects to boost creativity during interaction. Therefore, we defined two exemplary presets called *moods*. One preset is a calm setting with nice and warm sound, together with blue and violet colors, the other a sinister setting with dark and unpleasant sound supplemented by a red and black color scheme. The sands topography is colorized according to these color schemes by mapping the height value of each pixel of the height map provided by the Kinect onto a color ramp and projecting the resulting texture onto the sand. Using different color ramps, it is possible to create diverse visual impressions of the surface (see Fig. 3).

3.3 Simulation and Sound Synthesis

To achieve natural and naive to understand interaction, we decided to use a real-time fluid simulation that moves particles from left to right through the sands topography. Up to eight virtual AR sound points are placed onto the surface by the user. Each of these points attracts nearby particles and absorbs them, if they touch it. The rate of incoming particles of every point can be mapped onto an

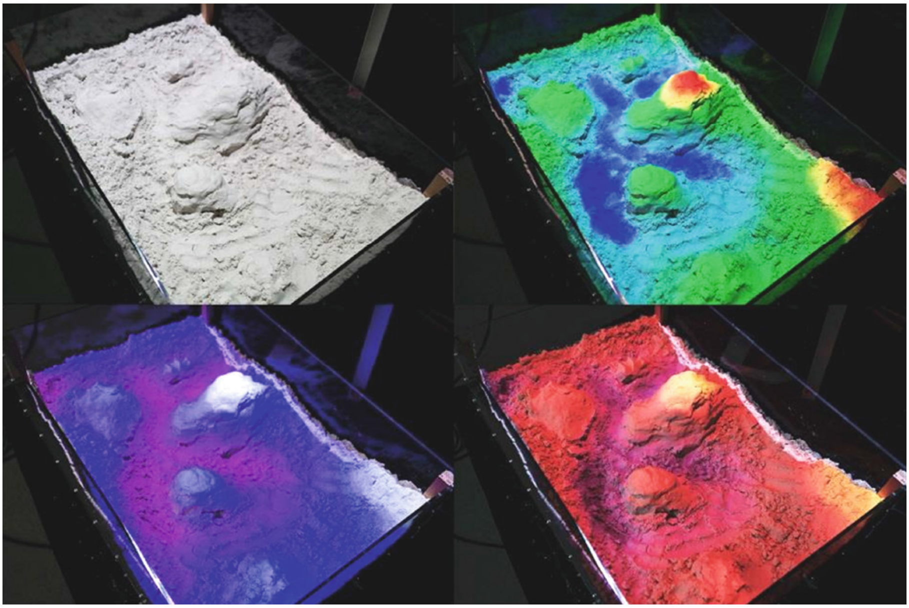

Fig. 3. Different visualization of height values using color ramps (Color figure online)

arbitrary sound parameter. In our system we chose to control the volume of predefined audio samples as it is an easy comprehensible way to interact with sound. To do this, the number of absorbed particles over a certain time is counted for every AR sound point, directly controlling the volume of the appropriate audio sample. We identified eight AR sound points and thus audio samples to offer an extensive amount of combination options for musical expression. The sound effects can easily be changed by replacing the audio samples (preferably using loop-able samples of a certain length). For advanced requirement we added OSC-support as a way to control other applications especially dedicated to creating music or sound via network. The decisions made focused mainly on the ease of use of the installation. The particle stream behaves in water-like manner that is easily understandable for the user.

We use a GPU-shader that simulates up to 8 million particles in real time and renders them onto a full-hd texture that is added to the background texture. To simulate real word behavior and thus predictable system reactions to user actions, we use a two dimensional Lattice Boltzmann simulation (LBM) in a multi-relaxation D2Q9-model to create water-like flows of particles (see Fig. 4). Hills and steep horizontal edges are detected in the Kinects depth image and added as obstacles in real time using an obstacle texture as shader input. LBM allows controlling the degree of turbulences by adjusting a single parameter [10] and can be implemented efficiently on the GPU. We selected a relaxation time value of $\tau = 1.98$ that creates both predictable and visual interesting flow effects. This value was selected through try-and-error.

The direction and velocity of the fluid is stored for each pixel in a fluid force texture. Additionally we create a same sized gradient map of the topography that describes the direction of the steepest slope for each point and drag particles towards this direction. Further, particles are attracted to AR-soundpoints with a gravity-like force. Each particle moving through the topography can look up the fluid and gradient force by accessing the stored pixel values at their position in 2D-space and calculate the attracting force of each soundpoint. All forces are then weighted and added in every frame to form a combined force that determines the momentary flow direction of each particle.

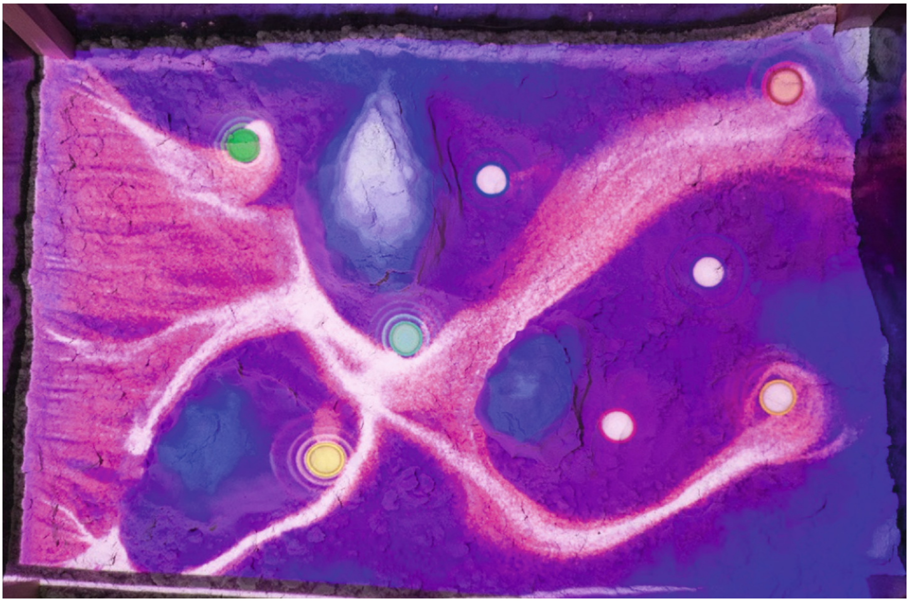

Fig. 4. Virtual particles flowing through the topography

3.4 Hand Tacking

For interacting with AR content we implemented a heuristic approach for hand tracking using depth data form the Kinect 2 sensor. We combine two different tracking methods: contour tracking and segment tracking.

Contour tracking is a popular approach to detect fingers in RBG-images. The background of a image is algorithmically removed so only pixels of the users hands are left. This can be easily performed in depth images as it is possible to remove all pixels that are farther away than a defined background threshold. The outline of the remaining pixels can be further examined to detect fingertips and other contour points of interest by their distinct features. In this project we use the radial distance method [13] to detect fingertips positions.

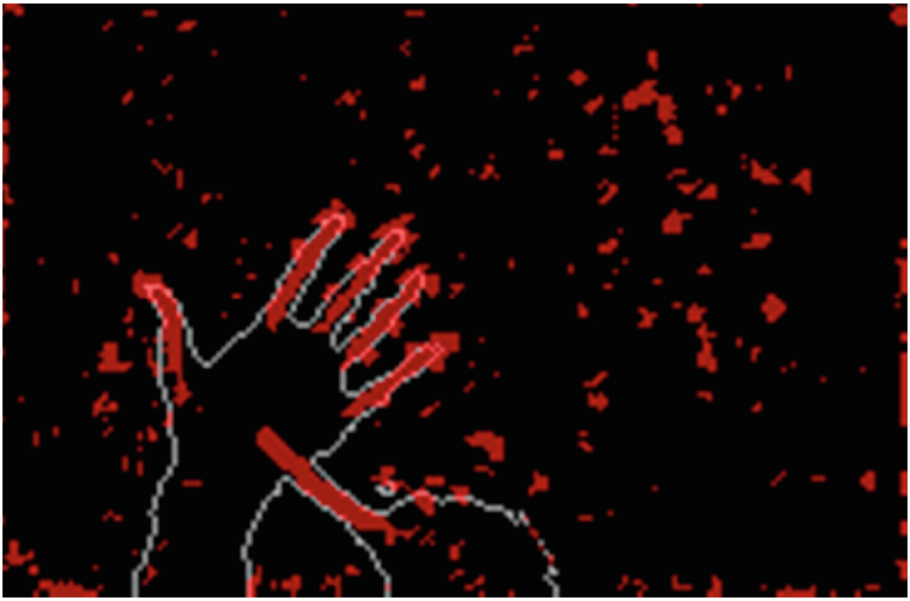

Fig. 5. Positive response (red) to fingerlike objects in the depth image (Color figure online)

During on-body interaction one hand masks the other so that contour tracking alone is not sufficient. In this case we detect finger-like objects in the image with a custom template matching algorithm that responds positive to pixels that are bounded by two opposing steep edges in vertical, horizontal or diagonal direction (see Fig. 5). A similar approach was presented in [3] for horizontal fingers. Positive responding pixels are then converted to a binary image which is segmented to identify connected positive responding pixels. If a segment's area size lies between a minimal and a maximal adjustable threshold it is considered a finger. We determine the distance between the lower positioned left hand and fingers of the right hand to detect touch events. If the distance of any finger lies below 15 mm, a value that is mostly defined by the Kinect sensors noise, a touch event is initiated at that position.

Our custom on-body hand tracking system was designed specifically for the presented setup and focused on resource efficiency to guarantee low-latency interaction and jerk-free visualization of the particle stream even on non state-of-the-art computers. It was not compared to known available hand tracking systems, as all requirements could be met.

3.5 Interaction Techniques

Interaction with the system is based on three different concepts:

First, it is possible to interact with the kinetic sand in a very natural way, like a child playing in a sand box. The user can sculpt the surface with both hands and watch the resulting changes of the particle flow. He or she can construct walls or hills that deflect the particles onto AR points or away from them. Forming holes causes particles to accumulate and digging ditches allows a precise control of particle streams.

Fig. 6. A user interacting with the menu projected onto his hand

Second, the user controls AR content and settings from the on-body menu which is projected onto the left hand if it's palm is facing upwards and it's thumb spread away (see Fig. 6). This pose is unlikely to occur during interaction, but easy to achieve deliberately, when the user wants to use the menu. To differentiate whether the system recognizes a left hand palm up or a right hand palm down, the user's position is tracked and estimated by the depth-sensor as well. For interaction we use a graphical menu projected onto the users hand controlled by touch and swipe events (see Fig. 6). Considering todays wide spread of mobile devices like smart phones and tablets, control by touch and swipe is presumably an easy to learn interaction technique for many potential users. Thus we implemented touch buttons and swipe gestures to scroll through a list of buttons as these are some of the most basic actions in mobile smart phones. Using this menu, the user can control essential features in our systems like adding new AR soundpoints onto the sand or remove existing points. Additionally, the user can switch between the two *moods* to change the produced sounds and production environment according to his or her needs.

Last, the user can perform gestural input. Depending on the operating mode, the user can delete or remove existing points. In order to do so, the user can select AR points by pointing a finger at them, just like indicating specific objects in the real world (see Fig. 7). The selection of the point is triggered after a short delay, to avoid the midas touch problem. Depending on the interaction mode, the selected point is than either removed from the sand surface or attached to the selecting finger to be placed somewhere else. The midas touch problem describes the difficulty of differentiating intended selection and thus positively provoking an event from unintended saccades that occur in human movement. Placing a AR point is handled in a similar way by pointing to the desired location. Further, the user can push particles with hand swipes in any direction or use his or her hand to temporarily block the particle flow.

Fig. 7. Several users interacting with and watching the exhibit

4 Evaluation

We carried out an informal user study during an exhibition of the system. We observed subjects during interaction with the system to get a first impression of how our system would affect user experience and how long users, who are interacting with the system for the first time, would need to comprehend the interaction techniques. The exhibition's participants were invited and had vastly divergent backgrounds. Our artistic installation was one of several exhibits, and users could approach it freely, interact autonomously or watch non-interactively (see Fig. 7). We observed both interaction with and without support by a developer.

4.1 Material and Methods

We used the sandbox described in this paper in a slightly darkened room during an exhibition of several mixed-reality and human-computer interaction research prototypes. All system components of our prototype, that were not relevant for interaction, e.g. computer screens, audio speakers or mouse and keyboard, were as hidden as possible to let subjects focus on the sandbox and projected content.

To convey the idea behind our system, we first gave some users a quick introduction to the system and design concepts by demonstration of the basic interaction techniques with the sand and the on-body projected menu through a researcher. Other users approached the system without mediation. After that, users were allowed to freely explore the interface by themselves and ask questions. We observed users during their interaction and noted down actions, statements and emotional reactions to generate qualitative data. During observation we tried to be as neutral as possible to avoid influencing the users opinions or exploring behavior. The observation of one subject ended as soon as a subject did not want to interact with the system anymore. Our main interest were user experiences during interaction and the usability of the system as a whole. The users did not know that they were observed, as the observer stood slightly away from the installation, but where enlightened later and shortly interviewed (if content with the use of the collected data).

4.2 Participants

About 50 subjects, mostly male, roughly ranging in age from 20 to 50, tested our system. As mentioned before, the subjects were not invited solely for the purpose of testing our exhibit. Rather, they were a random set of the exhibition's visitor entirety. All subjects were used to interacting with media content and computer systems, especially handheld mobile devices. Just a few subjects were familiar to the field of sound production or trained in using a musical instrument. In most cases subjects spent up to 15 min interacting with the system.

4.3 Observations

To most users our system had an inviting character. Forming the kinetic sand with bare hands provided an interesting haptic perception that invoked a positive reaction in most users. However, a few subjects refused do dig in the sand with their hands and preferred to use small shovels. Visitors were impressed by the possibilities of real time particle interaction and had fun to play around with particles. The coherence between landscape sculpting and particle flow was easily perceived by all participants. The on-body interaction menu was similarly well received, yet less intuitively learned. During interaction, by-standing visitors enjoyed watching the user's interaction with the system or tried to simultaneously use the system. Many subjects commented positively on the combination of sounds and suitable visuals and thereby induced atmosphere.

In many cases subjects focused on performing specific interaction techniques and not on exploring the possibility of sound production. In most cases, the first interaction of users was digging holes or trenches to direct particle flow and shortly after that, forming hills and walls to prevent particles from reaching specific sound points. The basic interaction concept of controlling the particle flow by forming the kinetic sand was very easy to understand and for all test subjects successfully applicable. Eventually, after firstly redirecting the particle stream to hit or avoid a specific sound point, users deliberately used gestural input and moved sound points into particle streams or out of them and listened to the changing soundscape. Users needed a small number of successful trial and error passes to figure out how to perform the selection process of virtual sound points correctly.

However, the on-body interaction menu was more difficult to comprehend. Some users needed a detailed instruction on the positioning of their hands. In many cases users recognized the analogy to mobile hand held devices and were able to learn the intended interaction techniques within several actions. Overall, users switched rarely between tangible input and selection gestures on one side and on-body interaction on the other and preferred to use different interaction techniques successively.

The used technologies and algorithms were fast and accurate enough to allow precise real-time interaction. The actual latency of the system was not measured, but no user commented on this matter negatively. All visitors used the system in the designated way, no matter if they were instructed about the system before or not. Furthermore, users tended to employ slow movements, reducing the effects of potential latency. However, we recognized this factor quite late in the observation, so we can only assume why this behavior occurred. One reason might the rather soothing atmosphere of the produced soundscape. The soundscape in general was rated more pleasant than disturbing. In some cases, user hands which were uncommonly large or small were not tracked well. This was caused by the scale-invariance of the system. Future approaches will employ a relative detection method. In almost all cases subjects needed some time to accommodate to the slower and less accurate tracking of touch events compared to touch screens but were eventually able to utilize the on-body interaction menu.

5 Discussion

Our interface is a new approach for interacting with music and sound using a continuous tangible user interface. Like other prototypes in this field, it shows that users have fun interacting with systems of this kind. Especially interactive projection mapping installations are well known tools for audience fascination (the examples stated in Sect. 2 are non-exhaustive). In our system, we use the data obtained to control further dimensions, and add the possibility to change system and interaction parameters through on-body projection. Even though the study was performed in a rather simple way with an early iteration of the prototype, some observation were made that influenced our future work.

In the last years some prototypes for on-body interaction and just very few prototypes for continuous tangible music interfaces have been developed. Our prototype shows that both research fields can be combined to add abstract features to continuous tangible user interfaces and can be implemented in an ready-to-used installation. Users need just a few successful actions to learn how to use tangible interfaces and just a few minutes to transfer interaction techniques from mobile device interaction to touch-event-based on-body interaction. Combining both seems like a good approach for system design, providing a vast multiplier to the interaction options, and should be further investigated by developing new systems and evaluating them in real world scenarios to identify possibilities and applications.

Fig. 8. Interactive indoor room-sized installation, augmenting the entrance area of the tanzhaus NRW

We made observations similar to Beckhaus et al. in [1]. Users are very interested in new forms of interacting with computer systems and show a natural curiosity about exploring new ways of interacting with digital content. By using well-known objects, for example a simple sand box, and adding further digital features, new systems emerge, that provide an inviting character to users. However, most prototype systems in this area only show feasibility and introduce possible applications; a formal study on user experience and system usability that confirms these statements in a quantitative way has not been conducted yet. Using statistical data, it would be possible to compare these new approaches of human-computer interaction to more established forms of interaction.

Our interface was mainly used in a laboratory for human computer interaction, and was not exposed to a real world scenario. Possible applications can be professional sound production, therapy and entertainment. To investigate these potentials our system can be further examined with both professional and non-professional users. From a production perspective, it would be very interesting to investigate how TUI-based music interface interaction compares to the common use of synthesizers in music production and if interface-dependent effects on the creativity of an artist are measurable.

6 Deployment

The system's engine was reused in a somewhat different yet still audio-visual installation, proving the feasibility of our basic concepts. The provided task was to create an interactive indoor room-sized installation, augmenting the entrance area of the tanzhaus NRW facility. The name of the installation was "Kinetic Stream" (see Fig. 8).

We decided to reduce the interaction dimensions of the presented system to solely the well perceived particle flow simulation including terrain detection for guidance of the particle stream. This decision was made after extracting the insight from our evaluation that this part of the system was easily understandable even without any explanation. Instead of sand, the visitors themselves were detected as "landscape" in a 4 m × 4 m wide area. Additionally, the particle stream was extended to flow down the wall first, like a waterfall. In this iteration of the system, the wall did not react to user interaction.

Representing obstacles for the system, visitors to the installation could block the particle stream by standing inside the tracking area or even lie on the floor. As there were no sound points present at this versions, the amount of persons inside the tracking area was counted. Next, each person was simulated as a sound point fully loaded by particles, regardless of the visitor's actual position in the stream (resulting in triggering the playback of an additional audio sample for each user in the tracking area). As visitors were to block the particle stream as if being elevated areas in the sand box, a coincident representation as particle absorbing sound points would have been contradictory. Furthermore, users quickly understood that they could change the soundscape simply by changing the number of persons inside the tracking area.

This variation of our system was very well perceived. The amount of positive feedback encouraged further development of this reduced version of our prior suggested installation, but only for larger scale installations. Therefore, in the latest iteration, we included the radar scanner radarTOUCH[1] to include touch-detection for the particle stream on the wall. Touching the wall simulates an

---

[1] http://www.lang-ag.com/de/produkte/touch-solutions/radartouch.html.

elevated area, deflecting the particle flow. The design of the particle flow was changed to resemble water, simulating a waterfall and small river (see Fig. 9). Inside the river, a small school of fish follows the user (for further entertainment). This extended version was exhibited during another exhibition at our lab, just like the one described in "Evaluation".

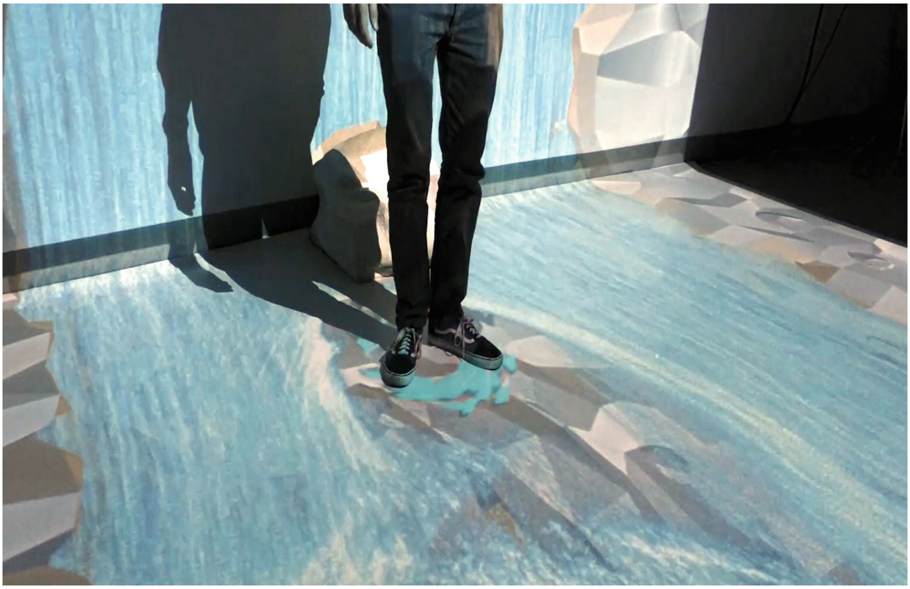

Fig. 9. A user deflects the water like particle stream by standing in the tracking area

We scraped the idea of implementing the on-hand interaction in this installation, despite the opportunities this would have featured. This is mostly due to the fact that we assumed inaccuracy for the hand detection, due to the necessary distance of the depth sensor to cover a 4 m × 4 m area and due to masking effects.

Even though this reduced version will be further developed and deployed, we still want to pose the question, if the originally intended combination of spatial and continuous tangible interaction with on-body projection and interaction can provide a method of interaction that is ready-to-go and easy to learn, yet offers a vast variety of interaction options.

7 Conclusion and Future Work

Our prototype system was successfully tested. Design concepts were well received and interaction techniques were easy to comprehend. The AR-sound sandbox and the underlying concept of real-time particle interaction is a new way for creating music and sound that can presumably be improved to be usable in real life situations. In our system, the user modified the volume of predefined audio samples, but trough mapping lots of different sound generators can be addressed (see Sect. 3.3). A reduced version of our system was deployed and successfully operated for several successive days, substantiating this presumption. In this version of our system, the users engaged more with the visuals, while audio was used to create an appropriate atmosphere.

Additionally, our project and evaluation shows that stationary on-body interaction systems can be used to add further control mechanics to sometimes restricted tangible interface interaction. However, in contrast to well designed tangible interaction that utilizes much more natural forms of interaction, users need to learn how to use graphical menus that are controlled by on-body interaction. In this context, touch event based interaction seems to be a good approach, as many users are familiar to smart phones and tablets today, which allows them to transfer interaction techniques from these devices to on-body interaction.

In future approaches we want to improve our algorithms to allow for quantitative evaluation using standardized measurements for user experience e.g. User Experience Questionnaire. The conducted study does not cover long-term use of our system and the setup was not ideal for quantitative evaluating, thus, important aspects for interaction were not investigated sufficiently yet. We further want to develop comparable interfaces that adopt the interaction techniques and investigate differences during user interaction with the systems. We currently plan to create a full TUI version that uses tangible artifacts instead of AR-soundpoints, and a digital version that is controlled in a more common way of human-computer interaction via mouse and keyboard to investigate differences in usability and interaction aesthetics of these approaches.

Acknowledgments. This work was partially supported by the Creative Europe EU Program (project The Peoples Smart Sculpture).

References

1. Beckhaus, S., Schroeder-Kroll, R., Berghoff, M.: Back to the sandbox-playful interaction with granules landscapes (2008)
2. Colter, A., Davivongsa, P., Haddad, D.D., Moore, H., Tice, B., Ishii, H.: SoundFORMS: manipulating sound through touch. In: CHI Extended Abstracts, pp. 2425–2430 (2016)
3. Harrison, C., Benko, H., Wilson, A.D.: OmniTouch: wearable multitouch interaction everywhere. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, pp. 441–450. ACM (2011)
4. Harrison, C., Ramamurthy, S., Hudson, S.E.: On-body interaction: armed and dangerous. In: Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction, pp. 69–76. ACM (2012)
5. Harrison, C., Tan, D., Morris, D.: Skinput: appropriating the body as an input surface. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 453–462. ACM (2010)

6. Jordà, S.: Sonigraphical instruments: from FMOL to the reacTable. In: Proceedings of the 2003 Conference on New Interfaces for Musical Expression, pp. 70–76 (2003)
7. Levin, G.: The table is the score: an augmented-reality interface for real-time, tangible, spectrographic performance. In: Proceedings of the International Computer Music Conference (ICMC 2006), New Orleans, USA (2006)
8. Mistry, P., Maes, P.: SixthSense: a wearable gestural interface. In: ACM SIGGRAPH ASIA 2009 Sketches, p. 11. ACM (2009)
9. Piper, B., Ratti, C., Ishii, H.: Illuminating clay: a 3-D tangible interface for landscape analysis. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 355–362 (2002)
10. Razzaghian, M., Pourtousi, M., Darus, A.N.: Simulation of flow in lid driven cavity by MRT and SRT. In: Thailand: International Conference on Mechanical and Robotics Engineering (2012)
11. The Green Cat Collective: Sand Noise Device An Augmented Reality Musical Sandbox. http://www.sandnoisedevice.com/
12. Wilson, A.D., Benko, H.: Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In: Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology, pp. 273–282. ACM (2010)
13. Yoruk, E., Konukoglu, E., Sankur, B., Darbon, J.: Shape-based hand recognition. IEEE Trans. Image Process. **15**(7), 1803–1815 (2006)