DESIGNING INTUITIVE SOUND INTERFACES: A HUMAN-CENTRED APPROACH COMBINING PAPER AND PHYSICAL CONTROLS

Jan-Torsten MILDE[1]

[1]*Computer Science and Digital Media*, **Fulda University of Applied Sciences**, Fulda, Germany

ABSTRACT

This paper presents a novel system for designing synthesizer interfaces, focusing on a human-centred approach. It explores the challenges of traditional synthesizer interfaces and proposes a flexible system using paper prototypes enhanced with 3D-printed controls and webcam tracking. The system allows for intuitive manipulation of sound parameters through physical interaction with these elements, offering a low-cost and adaptable method for creating both conventional and experimental musical instrument interfaces. The system's potential is illustrated through use cases involving the recreation of an analogue synthesizer interface and the control of a multi-channel auditory display. The research aims to bridge the gap between the technical complexity of sound synthesis and the creative expression of musicians.

1. INTRODUCTION

The development of synthesizers presents a unique challenge situated at the intersection of technology and design. A central aspect lies in the close interaction between the configuration of the technical sound engine and the design of an intuitive user interface. This interface is crucial for effectively supporting the sound design process and maximizing creative expression.

The sound engine, the core of the synthesizer, determines the sonic possibilities and the quality of the generated sounds. Simultaneously, the user interface must be designed to provide users with easy and intuitive access to these possibilities. A successful user interface enables musicians and sound designers to realize their sonic visions without unnecessary technical barriers.

In this paper, we examine the various aspects of this interaction.

2. THE DESIGN PROCESS

The design of a synthesizer interface is a complex balancing act that must meet a wide range of sometimes conflicting requirements. As the 'face' of the instrument, it should be aesthetically pleasing and ideally reflect the manufacturer's corporate identity. At the same time, intuitive usability is crucial, whereby the design can be based on established conventions to facilitate familiarization. A clear and logically structured interface is essential to maintain an overview of the numerous parameters and functions. However, the interface can also claim to reflect the underlying complexity of the sound engine. In this case, a deliberately challenging design is chosen that encourages the user to decipher the functionality of the synthesizer like a puzzle. This approach can lead to a deeper understanding of sound synthesis and innovative sound designs, but carries the risk of a steep learning curve and reduced accessibility for beginners.

The aim of this research is to develop a flexible and user-centered system for designing synthesizer interfaces that liberates the design process from conceptual constraints. We pursue an approach that combines the creative freedom of drawing-based methods, such as paper prototypes, with the integration of typical physical controls. The extensibility of the interaction elements should be conceptually unlimited to enable the development of novel instruments. The design process is understood as an iterative and collaborative process that provides for the early involvement of users and sound designers. The generated design artifacts are prepared in such a way that they enable seamless implementation into a digital sound engine. A central aspect is the coupling of physical interaction with the design prototypes to the sound engine, in order to be able to evaluate the interplay of interface and sound synthesis in an exploratory manner. By being open to new forms of interface design, the development of innovative and intuitive musical instruments is to be promoted.

2.1 Design challenges

The design of synthesizer interfaces is associated with a number of complex challenges that can significantly hinder the creative process. A central difficulty lies in the discrepancy between the synthesizer parameters and the resulting auditory output. This discrepancy leads to impaired usability and can be regarded as a problem of human-computer interaction (HCI). Users are forced to learn the domain language of the synthesizer, rather than communicating their creative ideas in an intuitive and needs-appropriate manner. The conceptual distance between the perceptual/semantic space and the parameter space of a synthesizer is large and complex, requiring extensive domain knowledge to translate between these spaces. (see [1–4]). More specifically, the following challenges arise:

- *Complexity of real sounds*: It is difficult to specify synthesis parameters in such a way that real sounds can be satisfactorily reproduced.

- *Conceptual distance*: Many musicians perceive a gap between their intuitive ideas of timbres and the control of synthesis parameters.

- *Navigation of conceptual spaces*: Users must navigate the parameter space, the perceptual space, and the semantic space when using a synthesizer.

- *Specification of timbres*: The core problem of synthesizer programming lies in the mapping between the perceptual and semantic space of timbre and the space of synthesizer parameters.

- *Technical terminology*: Parameters are defined in technical terms that relate to their relationship to the synthesis engine, making it difficult for users to relate them to their own perceptual/semantic conception of the associated audio output.

- *Steep learning curves*: Different synthesis methods require different background knowledge, and a significant time investment is required to learn the functionality of different forms of synthesis in depth.

3. HUMAN-CENTERED DESIGN

In the context of synthesizer interface design, the concept of Human-Centered Design (HCD, see [5, 6]) is gaining increasing importance. HCD describes a development approach for interactive systems that places users, their needs, and requirements at the center. By applying insights and techniques from the fields of human factors/ergonomics and usability, the goal is to design products, systems, or services that are user-friendly and useful. This approach aims to increase effectiveness and efficiency, improve user well-being and satisfaction, increase accessibility, and promote sustainable solutions, while minimizing potential negative impacts on health, safety, and performance.

HCD builds on participatory action research by utilizing the active involvement of users not only for documentation but for the collaborative development of solutions. In the early phases of the design process, methods such as immersion, observation, and contextual design are emphasized to gain a deep understanding of the problem and user needs. This is followed by phases of collaborative brainstorming, modeling, prototyping, and implementation.

In the field of synthesizer interface design, this means that development must originate from the needs and creative processes of musicians and sound designers, rather than from technical parameters. By integrating HCD principles, interfaces can be created that are intuitive, inspiring, and accessible, thus enabling optimal interaction between humans and machines. The evaluation of the success of such interfaces ideally occurs through methods of system usability measurement and by obtaining direct feedback from the user community.

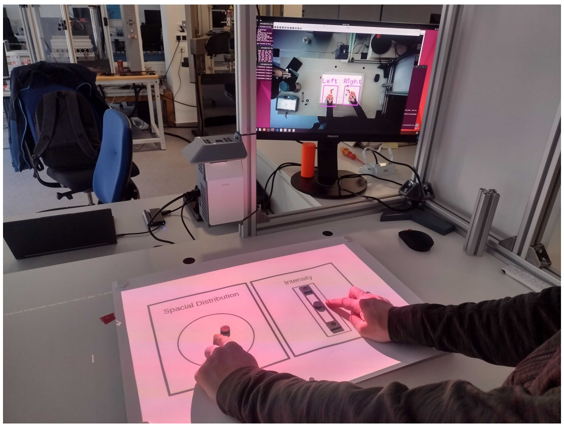

Figure 1. A minimal control for the auditory display: the circular control sets the position and distance of the sound source, the linear control is modifying timbre and volume. The graphical part of the interaces is projected onto the surface. The real time hand tracking calculates the parameter values.

3.1 Enhanced Paper Prototyping with 3D-Printed Controls

The integration of 3D-printed controls into paper prototypes significantly enhances the synthesizer design process. They enable more realistic representation, improved usability testing, faster iterations, and the creation of functional prototypes. The combination of a paper prototype with 3D-printed controls such as sliders, knobs, or switches offers a range of decisive advantages in the synthesizer design process. The approach will increase realism and facilitates haptic feedback:

- While paper prototypes represent the visual arrangement and layout of controls, 3D-printed controls enable realistic haptic feedback.

- Designers and users can experience the actual feel of knobs, switches, and faders, which is crucial for assessing usability and the user experience.

- The physical feel of the controls allows for a much more accurate assessment of ergonomics and ease of use.

The integration of 3D-printed controls significantly enhances *usability testing* by enabling a more realistic and meaningful evaluation of user interaction with synthesizer interfaces. By allowing users to interact with physical controls as they would with a real instrument, more precise insights into usability and potential operational errors are gained. This combination of visual representation and haptic feedback contributes to a more comprehensive and informed assessment of interaction design, which in turn supports the development of more user-friendly and intuitive synthesizer interfaces.

3D printing enables rapid and cost-effective creation of various shapes and sizes of control elements, facilitating accelerated iterations and adjustments within the design process. Designers can efficiently test diverse variations

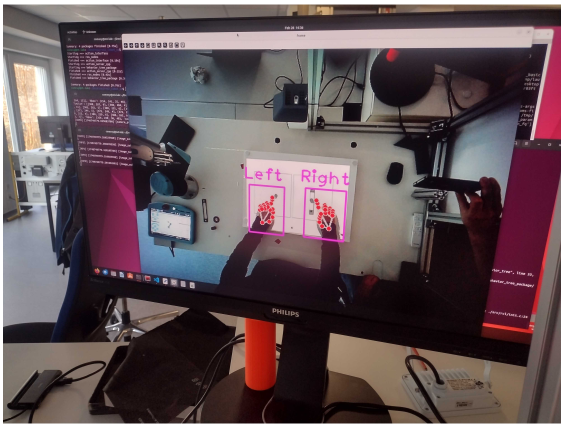

Figure 2. Tracking the user's hands in realtime: the camera based aproach is tracking the hand posture in real time. The UI elements are controlled with the tip index finges, allowing for two handed operations of the UI.

Figure 3. A set of 3D-printed controls: two circular controls with 4 markers each, a 2D-field with three markers, a 2D-field with a matrix inlay allowing to set discrete values, two sliders allowing to set a continous value or setting an interval. ArUco (see [7]) patterns are used to identify the controls.

of knobs, switches, and faders, selecting the optimal solution to meet their specific needs. This capability allows for swift implementation of modifications and customizations, significantly expediting the overall iterative design cycle. In addition, the physical controls can be employed to enhance the visual aesthetics of the synthesizer, allowing designers to utilize various materials and surface textures to achieve an appealing and high-quality appearance. Furthermore, an enhanced paper prototype incorporating 3D-printed controls proves highly effective for presentations and demonstrations, providing a tangible and visually compelling representation of the design.

4. GENERIC 2D INTERACTION ELEMENTS

In our approach, the design of interactive interfaces for synthesizers is realized by implementing 2D elements defining a rectangular region with a defined reference point.

These regions enable the intuitive control of synthesizer parameters by capturing x/y vectors and rotation angles, relative to the origin, centroid, or a user-defined rotation point. Color-coding the physical markers can provide visual cues about the current status or parameter value.

Interaction occurs through touching and occluding the physical markers with a fingertip of the index finger, triggering an action. The system only reacts to the index fingers, ignoring occlusions with other parts of the hands. The subsequent change in the parameter value is mapped in real-time by the movement of the fingertip (see fig. 2). Once the physical marker becomes visible again, the action ends. This method offers precise and dynamic control, enhancing the haptic experience of synthesizer use.

4.1 The design system

The design system consists of a metallic plate on which both the paper for graphical sketching of the user interface and the 3D-printed interaction elements are placed. Magnets are used to fix the paper and 3D-printed elements. Alternatively or in addition to the paper, an image can also be projected onto the surface. Interaction with the UI being developed is captured via a webcam positioned above the metallic plate. Image data processing is performed by a computer, in the simplest case using a Raspberry Pi.

A set of 3D-printed UI elements have been created:

- a circular control
- 2D control field
- discrete matrix inlay for the 2D control field
- slider

Each of the UI elements is holding physical markers. These markers are to be moved by the fingertips of the user. The UI elements is constraining the movement of the marker (see fig. 3).

The system software captures the positions of the user's fingertips and converts them into corresponding parameter values. This conversion is dependent on the specific physical interaction element which the user is interacting with. The determined values are then sent via the OSC (Open Sound Control) protocol to the sound engine, where they can be further processed and used to control sound generation (or anything else).

This system provides an exceptional flexibility and user-friendliness. There are essentially no limitations regarding the shape of the control elements, allowing for the rapid and intuitive implementation of virtually any conceivable interface.

In addition to familiar UI elements, strongly semantically oriented or primarily aesthetically designed elements can also be employed.

The system enables new elements to be learned in a very short time. To do this, the user only needs to 'traverse' all extreme values in the interface once, i.e., fully left, right, up, and down. This process is performed only once per element and requires minimal time. The configuration is then stored and the UI Element is fully functional.

A further advantage of the system is its low cost. For the prototype, a metal plate was procured from a hardware store, and the 3D-printed elements (2D fields, matrix inlays, circular controls, slider controls, and markers of various sizes) weigh less than 200 grams. This brings the material costs to approximately 15 euros. The webcam used is a standard model, resulting in total costs of approximately 80 euros.

5. TWO USE CASES

During the conception and implementation of the design system, two use cases were addressed. First, for the digital reproduction of an analog synthesizer (TheWasp, in German DieWespe), the 3D elements were developed step by step, and the software was implemented and tested. In a second step, the approach was transferred to multichannel auditory display and used by students as part of a course.

5.1 DieWespe

As a first example of the design process serves the development of a module for the virtual modular synthesizer VCV, which is based on the well-known British synthesizer 'The Wasp' by EDP. This synthesizer is a typical representative of synthesizers from the late 1960s and early 1970s and is based on classic subtractive synthesis.

The layout and design of the user interface follows the conventional structure for synthesizers, aligning layout with the signal flow of the underlying sound engine from VCO (or DCO) to VCF to VCA. Modulators and envelope generators are positioned horizontally below the unit to be modulated.

All sound engine parameters are controlled via continuous rotary knobs, with clockwise rotation adjusting from minimum to maximum value. Discrete rotary knobs function as (toggle) switches. The signal flow and strength of the modulators can be adjusted, and a touch-sensitive keyboard controls the fundamental frequency in equal temperament.

Sound design is conducted in a sequential manner, beginning with waveform selection, followed by filter parameter adjustment, and concluding with volume envelope manipulation.

The module's interface is designed to be read from left to right and top to bottom. This arrangement reflects a distinctly euro-centric user experience (and is also assuming that the standard user will primarily make adjustments with their right hand).

The user interface of DieWespe is created using the Inkscape graphics software. Instead of fully replicating every knob of the original interface, a smaller number of controls are defined. Each physical interface element acts as a meta control, therefore changing multiple sound parameters in parallel, a functionality, that is not available in the original device. We use a projector to display the interface onto the metal plate of the design system. The physical controls are placed on top of the projection (see fig. 4).

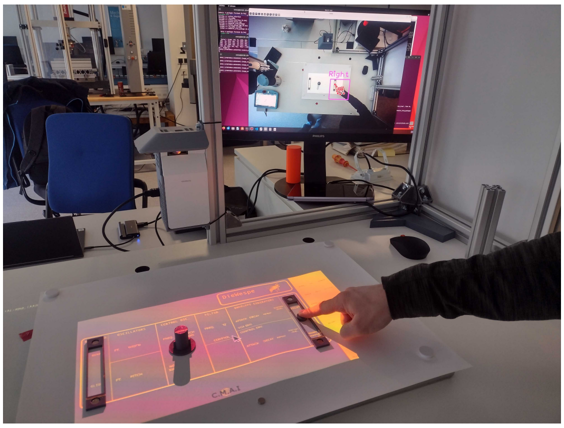

Figure 4. The UI of DieWespe: the Interface of a VCV Rack module is created using the inkscape drawing software. In this use case, we use a projector, allowing to project the interface onto the table top. The physical controls are placed onto the projection.

Figure 5. The auditory display: in this circular configuration, 8 speakers are surrounding the control panel. In addition the multi-media studio provides dynamic lighting and a green screen to visually extend the sonic experience.

5.2 Multichannel Auditory Display

Inspired by a technical setup observed at SMC 2021 in St. Etienne, a mobile 16-channel auditory display system had been set up as part of the multi-media studio of the faculty (see fig. 5). This system comprises a mobile rack housing two optically coupled 8-channel audio interfaces and 16 active loudspeakers with height-adjustable stands. The setup is utilized both within and outside the university, serving as a teaching system and simultaneously functioning as a temporary exhibit in exhibition contexts.

Students are provided with the opportunity to experiment with interactive, spatial audio and music installations. Thanks to the modular design of the technical setup, diverse configurations and application scenarios can be explored. In the experimental setup, audio synthesis is performed using SuperCollider or MAX/MSP. The 16 loudspeakers are placed concentrically in two circles with diameters of 5 meters and 8 meters. Within the in-

Figure 6. A more elaborate control for the multichanel audio display: 3 sound sources can be placed in the 2d plane, discrete sonic features can be set using the matrix control, two circular control for each of the 8 channels, the sliders set volume and filter values.

ner circle, an interaction station is positioned, designed as a movable desk on which the control elements are arranged. The camera-based hand trackings carried out using a Raspberry Pi, which is connected via Wi-Fi to the separately located sound engine. The multi-media studio offers additional media possibilities, such as dynamic lighting (DMX), greenscreen, external camera tracking to be integrated into the sonic experience.

Two versions of a user interface for the auditory display have been explored. A minimal control (see fig. 1) for the auditory display consists of only two control elements: a circular control sets the position and distance of the sound source, while a linear control is modifying timbre and volume. The minimalistic graphical part of the interface was designed in a graphics program and is projected onto the empty white surface of the design system.

The more eloborate UI version (6) uses 4 control elements in varying configuration. The 2D control element allows to place 3 sound sources in the 2d plane. A number of discrete sonic features can be set using the matrix control, and with two circular controls parameters of the 8 output channels are modified. Finally the sliders set volume and filter values.

6. CONCLUSIONS

The design of synthesizer user interfaces presents significant challenges due to the disconnect between complex technical parameters and the intuitive needs of users. A Human-Centered Design methodology is proposed to bridge this gap, emphasizing user-centric development for enhanced usability. To this end, a novel and cost-effective system is introduced. This system integrates traditional paper prototyping with 3D-printed physical controls and webcam-based capture of 2D interactions. The primary goals of this system are to provide realistic haptic feedback during usability testing, accelerate the design iteration process, and enable the creation of a wide range of intuitive and accessible and fully functional user interfaces.

7. REFERENCES

[1] J. Shier, "The synthesizer programming problem: improving the usability of sound synthesizers," Ph.D. dissertation, 2021.

[2] A. Seago, "Analysis of the synthesizer user interface: cognitive walkthrough and user tests," *Department of Computing, The Open University*, 2004.

[3] ——, "Electronic sound synthesizer interface design," *Department of Computing, The Open University*, 2004.

[4] P. Cook, "2001: Principles for designing computer music controllers," *A NIME Reader: Fifteen years of new interfaces for musical expression*, pp. 1–13, 2017.

[5] G. A. Boy, *The handbook of human-machine interaction: a human-centered design approach*. CRC Press, 2017.

[6] M. Cooley, "Human-centered design," *Information design*, pp. 59–81, 2000.

[7] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, 2014.

[8] A. Seago, S. Holland, and P. Mulholland, "Synthesizer user interface design-lessons learned from a heuristic review," 2004.

[9] J. Kleimola, "Design and implementation of a software sound synthesizer," Ph.D. dissertation, MS thesis, TKK Helsinki University of Technology, Espoo, Finland, 2005, http ···, 2005.

[10] J. W. Ruffner and G. W. Coker, "A comparative evaluation of the electronic keyboard synthesizer user interface," in *Proceedings of the Human Factors Society Annual Meeting*, vol. 34, no. 5. SAGE Publications Sage CA: Los Angeles, CA, 1990, pp. 477–481.

[11] S. Potidis and T. Spyrou, "Spyractable: a tangible user interface modular synthesizer," in *Human-Computer Interaction. Advanced Interaction Modalities and Techniques: 16th International Conference, HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part II 16*. Springer, 2014, pp. 600–611.

[12] C. Rasmussen, "Evaluating the usability of software synthesizers: An analysis and first approach," Ph.D. dissertation, University of Guelph, 2018.

[13] C. Roberts, G. Wakefield, and M. Wright, "2013: the web browser as synthesizer and interface," *A NIME Reader: Fifteen Years of New Interfaces for Musical Expression*, pp. 433–450, 2017.

[14] M. Vail, *The synthesizer: a comprehensive guide to understanding, programming, playing, and recording the ultimate electronic music instrument*. Oxford University Press, 2014.

[15] C. Deckert, "Good vibrations from electronic circuits: The technological development of the synthesizer," 2024.