Hochschule für Technik und
Wirtschaft Dresden
University of Applied Sciences

# 25th Bilateral Student Workshop
# CTU Prague - HTW Dresden
# User Interfaces & Visualization

Dresden (Germany) 2022

Editors: Hans-Joachim Böhme
Miroslav Macík
Pavel Slavík
Dietrich Kammer
Jan Brose

# 25th Bilateral Student Workshop CTU Prague - HTW Dresden
# User Interfaces & Visualization

Hochschule für Technik und Wirtschaft Dresden
University of Applied Sciences

## Contents

# Multiprojection of Langweil´s model

Barbora Koudelková

Department of Computer Graphics and Interaction, Czech Technical University in Prague, Faculty of Electrical Engineering

**Abstract.** This work deals with the possibilities of applying multiprojection to the digitized Langweil's model of Prague. Examining the methods to enable the viewer a pleasant experience, the goal of this work is to design optimal procedures for rendering and displaying Langweil's model using 5 screens simultaneously, both for educational purposes (using the remote controller app developed by CTU for the presenter) and for random visitors of the museum – It contains extensive research, from VR prototyping to testing with real users. The result of the work is the creation of a video sequence according to the scenario of the Prague City Museum, prepared for future projection as a part of the museum exhibition.

## 1 Multiprojection and Langweil's model

Langweil's model of Prague is a model of the historical center of Prague made out completely of cardboard paper. The model was created during the years 1826 to 1837, by a servant of the Prague University Library, Antonín Langweil [1]. The real model occupies roughly 20 $m^2$ and must be stored in a glass display case, preventing further decomposition. During the years 2006 to 2008, the entire model was transferred to a digital form by Visual Connection s.r.o. in cooperation with DCGI, CTU FEE.

It was decided to develop ten chapters that would be shown and screened at the upcoming exhibition, based on the scenario presented by The Prague City Museum, using the digitized version of the model. Individual chapters will be part of the educational program for schools, regular visitors should be played a loop of these videos, although they are not consecutive. The projection should take place in a previously designed construction on five screens, four placed diagonally above the heads of the visitors, one, square in size, placed in the middle as a central table.

Projection onto a square table with a ratio of 1:1 and projection onto four beveled walls with a ratio of 7:2 was expected. The whole object is supposed to resemble a cube, or a house made out of a cardboard puzzle, and is complemented by a central projection table, as shown in figure 1.

## 2 Task execution

Over-detail capture was set to 1 mm in the geometric area (vertices, edges, outlines). The building models are partly flat and contain extremely detailed

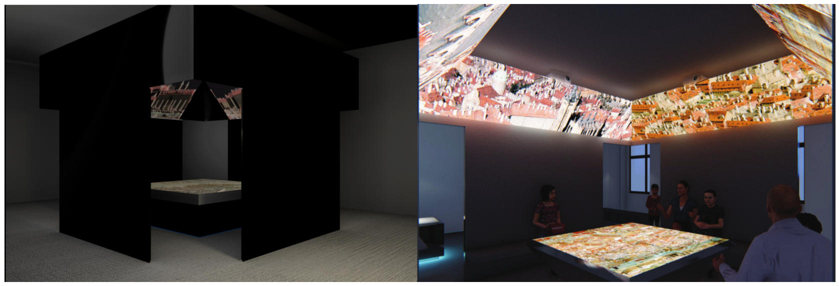


**Fig. 1.** Visualisation of real projection construction

textures. Due to the complexity of the model and a large number of external references to those textures, obtaining the image data was quite a challenge, and interventions in the final form had to be minimal. For this reason, it was necessary to constantly consult with the museum on how to sensitively capture the essence of the model. After many iterations, the museum agreed that, in general, it was not possible to use special effects for storytelling, animate individual parts of the model, or add some artificial objects to the scene. Furthermore, it was necessary to solve how content could be projected onto which screen. Many prototypes have been created, from physical ones to a complete mockup room in virtual reality. It was found during testing, that it is not possible to project videos asynchronously, with different content on every screen, as it caused dizziness. At the same time, it was found that videos could not be played on both the upper and lower central screens, for reasons already mentioned. The resulting renders with added voiceover were then deployed into the mobile application of the 5D projection system (developed by the CTU), which served as a remote controller for the lecturer during the educational program.

## 3 User testing

User testing took place on the grounds of the Ctěnice castle complex between April 4th to 7th, 2022. The pilot testing took place in cooperation with the lecturers of the Prague City Museum and has been tightly linked with an educational program thematically touching on individual chapters. The target group was students from the second grades of elementary schools and eight-year grammar schools. The testing program was as follows:

1. Playback of the complete video sequence (in construction mentioned in Chapter 1)
2. Introduction to the topic, distribution of partial tasks
3. Work in groups (outside construction)
4. Possible projection of other materials and historical sources
5. Evaluation of lecturer programs (guided interview)
6. Evaluation of the introductory video projection (questionnaire)

Unfortunately, the research was plagued by several errors. As can be seen from the course of the program in beginning of Chapter 3, all videos were played in sequence (see point 3.1), without any introduction to the topic, and therefore these videos were not worked on according to the assumptions stated in Chapter 1 for the lecturer program.

The screening took place in the construction mentioned in Chapter 1, with one main difference. In the original visualizations, the central projection screen was located approximately 65 cm above the ground, which was also taken into account by every prototype of multi-projection created. However, in the construction built for the user testing, the canvas was laid on the ground, which was causing troubles specifically with selective attention, where attendees failed to notice the switching between central and upward screens, because the screens were not in one line of sight, as was expected (see figure 2).

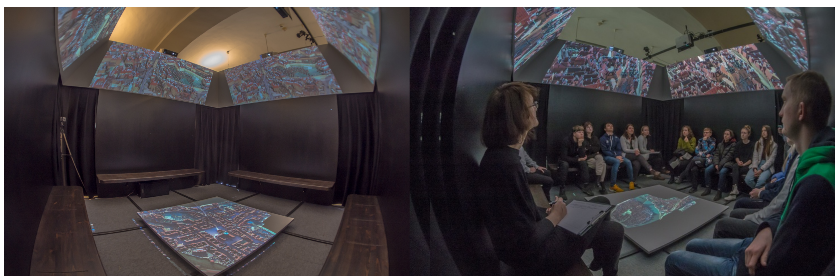


**Fig. 2.** Projecting in the real construction

This concrete deficiency was very often negatively depicted in the subsequent evaluation. The survey was carried out by a combination of controlled group interviews and anonymous questionnaires with both closed and open questions, in which 162 respondents took part.

From the conclusions of the statistical investigation, it is evident that the method of projection onto the central table and four walls are not well resolved.

Judging by the subjective feeling and the test results, the main problem is the expectation of additional content after the projection of the maps, while the user is not sufficiently alerted to the beginning of the content projected on the sides of the structure, and for this reason, always misses part of the upper projection. This shortcoming can be solved either by reducing the brightness of the central map after the initial projection or by completely darkening the lower projection during the upper screening, possibly by inserting an element that prompts the visitor to change his view to another projection surface. It was also obvious that animating certain elements on central map locations during the overhead projection was highly distracting and should not be used in future iterations of the work. Furthermore, the author does not consider it a bad idea to soundproof the projection space, as during the physical testing there were spatial reflections from the laminate boards of the structure.

At the same time, it is necessary to mention that, on the one hand, with regard to the rotation of projections, on the other hand, from the point of view of the bad verbal evaluation of the respondents, it would be necessary to change the concept, the length of the projection, and the content level of the scenario. Vast majority of visitors had the feeling, due to the amount and speed of information transfer, that they did not take away any valuable information from the projection or were constantly missing key points.

## 4 Conclusion

As already mentioned, in Chapter 2 and Chapter 3, it is not possible to project on all screens at the same time. It can also be stated that asynchronous screening often induces nausea and it must be taken into account when creating the story [2]. Another issue that needs to be addressed is the resolution of screen switching. At the present moment, visual clues are given to the observer at each switch, alerting him to the change in the location where the projection is taking place. At the same time, it was found that the scenario, targeted for educational purposes, is not well understood by the average visitor.

The museum, therefore, decided to create a continuous video sequence and collaborate with the art director on the script and story. For that reason, the work continues to be redone and will be part of the museum exhibition in the fall of 2023.

## Acknowledgement

My thanks go both to Ing. David Sedláček, Ph.D., and Prof. Ing. Jiří Žára, CSc., for their professional guidance during the creation of the work. I also thank The Prague City Museum, namely PhDr. Kateřina Bečková, the curator of Langweil's model, for her expert supervision.

This work has been supported by the Ministry of Education, Youth and Sports of the Czech Republic, grant No. CZ.02.3.68/0.0/0.0/18_067/0012327 (The Langweil model of Prague as a school historical source), and (partially) by the Grant Agency of the Czech Technical University in Prague, grant No. SGS22/173/OHK3/3T/13 (Research of Modern Computer Graphics Methods 2022-2024).

## References

1. Bečková, K., Fokt, M.: Svědectví Langweilova modelu Prahy. Praha: Schola ludus-Pragensia, ISBN 80-900-6688-7, 1 edn. (1996)
2. Johnston, W.A., Dark, V.J.: Selective attention. Annual review of psychology 37(1), 43–75 (1986)

# Design of an assistant for persons interested in study at CTU FEE

Adam Shorný

CTU - Faculty of electrical engineering

**Abstract.** This thesis deals with the problematic of students' transition from high school to university. The aim of this work is to find out whether chatbot technology could be useful in addressing this issue and how best to utilize it. The work includes an evaluation of the benefits of such a solution for high school students and the Faculty of Electrical Engineering. The work is mainly based on a survey of high school students. The implemented solution is based on the IBM Watson Assistant platform.

## 1 Motivation

There are many ways in which universities present themselves to convey the most information to potential students. However, it can be difficult to determine what information is really needed, and sometimes it can be hard to find.

Chatbot technology, which is based on questions and answers, could make this process easier. It will provide prospective students with the opportunity to ask about what they are interested in and receive clear answers. Thanks to the popular chat format, this form of information communication can impress the target audience and create a positive impression of the school as a whole.

Thus, the use of chatbot technology could bring several benefits not only for the prospective students, but also for the Faculty of Electrical Engineering. It offers another way of offering information to future students, which could result in increased interest among graduating students in future studies at CTU FEL or the faculty itself.

## 2 Collection of information

The aim of this work is to gather and consolidate the information that high school students need to know when transitioning to university. Based on this information, a chatbot can then be created to answer questions that high school students have about the Faculty of Electrical Engineering at CTU.

In the introductory chapter, I would like to present the procedure and method used to collect information for the preparation of the thesis. The data collection took place for approximately one month, specifically at the end of November and beginning of December 2020, when high school students start choosing a college. The information was obtained from three sources. The first method was a questionnaire given to high school students, mostly seniors, asking several basic questions. The second source was through interviews with selected high school students. The third and crucial source of information was an interview with an educational counselor at the Pierre de Coubertin Gymnasium in Tabor.

Obtaining all the information primarily served to implement the chatbot. Another benefit of the survey was obtaining contacts with high school students, who all promised future cooperation in the development of the chatbot. This collaboration mainly involved testing and providing feedback on the final solution. The same benefit was obtained from the contact with the educational counselor, from which validation of the key areas took place.

# 3 Evaluation of the questionnaire for secondary school students

This section describes the evaluation of a questionnaire that was used to gather information for the development of a thesis focused on using chatbots to help secondary school students. The questionnaire was distributed to 250 students from dozens of high schools and was primarily completed by final year students. The information gathered from the questionnaire and subsequent interviews provided insight into the perspectives of high school students on technical majors in college and the communication methods used by colleges to convey information to prospective students.

The findings showed that colleges have a communication problem with prospective students and that the main method of communication is through websites. The use of chatbots in this area could help to solve this problem by making it easier for students to quickly find the information they need through a question and answer system. The use of a chatbot could also bring greater attractiveness to the Faculty of Electrical Engineering and increase the interest of applicants in studying there. By communicating information through a chat window, the Faculty of Electrical Engineering would be closer to its target audience and provide another alternative to traditional communication methods such as websites.

# 4 Basic information about the Faculty of Electrical Engineering CTU

The chapter summarizes basic information about the CTU Faculty of Electrical Engineering (FEE). You are interested in potential applicants. Information is provided on the location of the buildings of the FEE, accommodation options for students offered by the Administration of Special Purpose facilities, study programs offered by the FEE and a simplified principle of admission to study programs. The chapter concludes with a summary of some other concepts that are different from those of secondary schools and thus important for applicants to clarify their meaning. The chapter is again based on the initial survey and deals with the following topics that are of interest to applicants seeking information about them.

The information in this chapter is based on an initial survey of high school students and consultation with the guidance counselor. These are basic information of practical interest to every applicant to the FEE. These topics are precisely the basis for the development of a chatbot to facilitate the acquisition of information about CTU FEE. The chatbot is designed to answer clearly and concisely most queries of a similar nature. The answers are not completely detailed, but contain the most important aspects that the user needs to know.

## 5 Chatbot Technology

The basic information about the chatbot technology is provided, followed by some examples of its current usage and the reasons behind our choice to use it for resolving information retrieval issues in the field of colleges.

## 6 Creation of the Chatbot Felák

The process of creating the chatbot, Felák, for relaying information to prospective students at the FEE, was divided into several sequential steps. Each step involved initial design, development, and testing or discussion of results. The implementation process was based on the findings from the initial survey to identify the most frequently asked questions and areas of concern for high school students. The style of responses was also influenced by the information collected about the presentation of colleges.

In the figure 1, we can see the mindmap of the chatbot. There is the beginning of each conversation - 'Hello!!!' and all topics, which he can answer divided into logical areas.

## 7 Assessment of the current state of the chatbot Felák

This chapter provides an overview of the current status of the Felák chatbot and its perception by various user groups. The evaluation was conducted using several methods, including user testing and feedback through questionnaires, consultations with the guidance counselor at Tabor High School, and discussions with the public relations team at the Faculty of Electrical Engineering.

The chatbot technology is highly sophisticated and has the potential for continuous improvement. Future development could include adding information about sports and activities at the faculty, general information about teachers, scholarships, and clarification of terms like "dean" and "rector."

To increase the chatbot's accessibility, it may be necessary to provide translations into English and Russian, especially considering that not all students at the Faculty of Electrical Engineering are from the Czech Republic. Additionally, providing information specific to international students, such as recognition of prior studies and tuition fees, could be valuable.

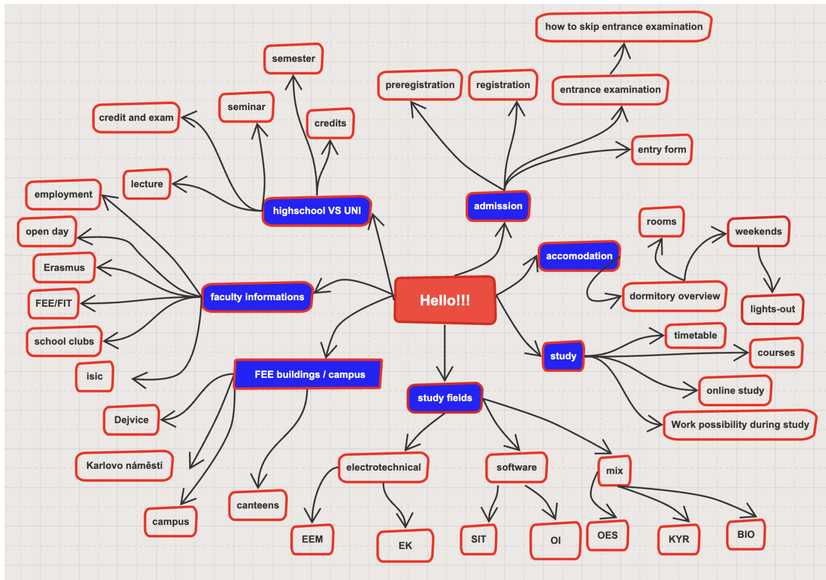



**Fig. 1.** Mindmap of chatbot

To fully implement the chatbot, it would be necessary to collaborate with the public relations department at the Faculty of Electrical Engineering. This would involve acquiring a license for the chatbot's use, deciding on its deployment method (e.g. as a standalone website or integrated into the existing faculty website), and maintaining its accuracy through regular updates.

# Sonification of a juggling performance

Vojtech Leischner (0000-0001-8399-3543), Jiří Balhar (0000-0002-2298-6304)

Czech Technical University in Prague

**Abstract.** We aim to sonify juggling performances with balls equipped with accelerometers, gyroscopes, and WiFi. The system measures acceleration during juggling and can distinguish when the ball is in the air, caught, or static. We also track the individual rotation of the ball in 3D. Therefore we can map time in the air to audio pitch and animate the virtual audio source position. Furthermore, we can pair each ball with a distinct musical instrument. The result is a movement-driven musical instrument. As the juggling patterns have inherent rhythm, we can use it to create a musical composition in real-time. Furthermore, we suggest other use cases for the system, such as virtual reality controllers, performance feedback, tracking or a MIDI interface.

## 1 Use Case

Besides a musical instrument, we propose implementing it as a VR controller [4]. We can map the rotations of the unit to the rotations of the 360 equirectangular projection on the monitor. Users can navigate inside the 360 photos by simply rotating the device.

Another valid use case might be a relative motion capture system or performance tracking. We can track time in the air and compare these readings with each other to help achieve user uniformity in his/hers throws. Similarly, we can incorporate the unit inside a football or baseball ball.

## 2 State of the Art

Till Bovermann et.al. created a similar setup for juggling sonification using multiple cameras and markers stuck to the juggling clubs [1]. Our proposed system is based on relative motion measurement rather than tracking absolute position in space. The advantage of using cameras is that we can clearly define the relationships of props to each other. The limitation is mainly the required hardware and calibration setup that hinders the system's usability as a controlled environment is needed. Using the IMU sensors and WiFi microcontroller directly inside the balls, we can use the system independently of the current environment in all lighting conditions.

## 3 Data flow

Sensors are connected to the microcontroller with integrated WiFi. We send UDP packets over WiFi containing the acceleration vectors and rotation as quaternions. We receive the packets on a computer or mobile device that analyzes the data and sends out events such as catch and throw with time stamps over OSC and MIDI. Users receive events in their preferred software, such as Ableton, Cubase, Reaper, and other DAWs. We have also included our custom sound engine sampler to showcase the possibilities without additional DAW software so the setup can be tested without it.

## 4 Graphical User Interface

We have implemented 3D visualization of rotation on cubes. Furthermore, we showcase real-time graphs showing acceleration magnitude in time and individual UDP packet receiving times, see Figure 1. On the left side of the window, we feature interactive control over the integrated audio engine. Users can choose to disable it when using other sound engines. We added toggle controls to individual parameters, such as mapping pitch to the spent time in the air or enabling respective audio sample banks.

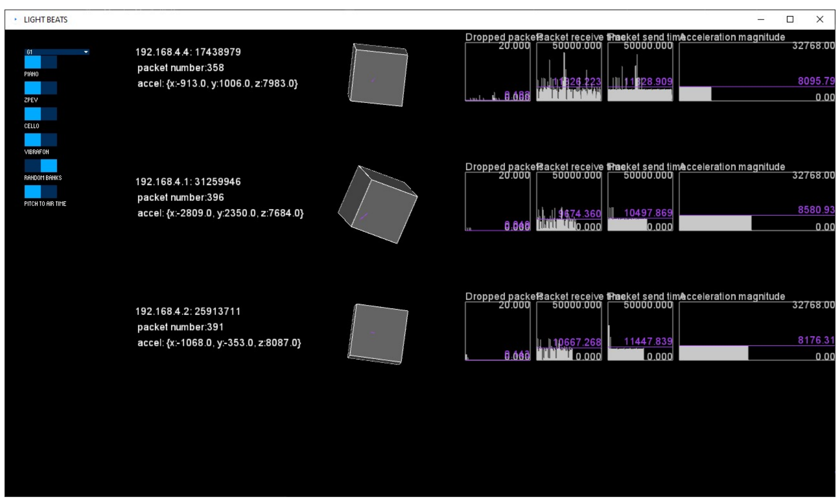



**Fig. 1.** GUI application running on a computer

## 5 Hardware prototype

We created the custom hardware for the prototype, see Figure 2. However, you can also buy finished modules with integrated sensors. The suitable modules are LILIGO® TTGO TAudio V1.6 or M5stick with needed sensors on board. We use a widely available ESP8266 microcontroller and MPU6050 IMU unit on a custom PCB board. We attached a 102535 Li-Po 800mAh battery to power the unit with a step-up and charging module TP4056. We also add LEDs for signaling the status of the device and absolute tracking with cameras if the user chooses to.

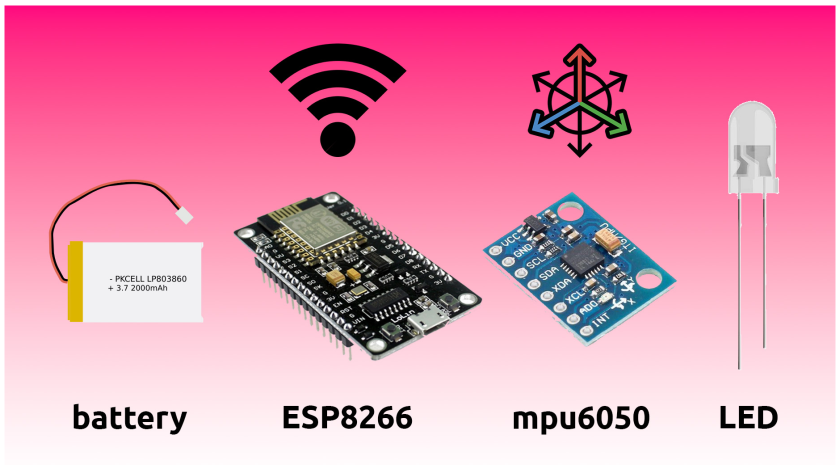



**Fig. 2.** Hardware prototype - battery powered WiFi microcontroller with IMU

## 6 Mapping audio to movement

The system's main objective is to provide a new expressive musical instrument that maps movement to the sound. We have come up with several strategies to achieve that. First, we create discrete events or triggers when the ball is thrown or caught in the hand. We can map these events to MIDI note-on and MIDI note-off events. Like this, we can have a note being played while the ball is in the air. The only problem is that we need to know the trajectory of the ball at the time when it is thrown. That is why we have finally decided to play the sound when the ball is caught. That way, we can map the time spent in the air to the length of the note being played and its velocity. Besides these discrete MIDI and more general OSC events, we have also experimented with the integrated audio sampler.

We use the same strategy with the audio sampler as with MIDI notes. We play the sample when the ball is caught. Furthermore, we also use semi-random sample variations - the octave is chosen randomly from the predefined range to achieve more variety in sound composition. Each ball is given its own sound sample bank to play. For example, we mapped one ball to the sounds of violoncello, the second one to vocals, the third to the piano, etc.

## 7 Future research

We want to extend the system by connecting to our previous research on spatial audio web player [3]. We can map the movement in the air to spatial audio panning parameters. Another approach we would like to explore is using multiple speakers where each speaker would be mapped to an individual ball. That way audience can more easily recognize the interactive relationship between juggling balls and the music produced. Furthermore, we want to move beyond the simple audio sampler and integrate granular audio synthesis to allow us finer control of the sounds.

## 8 Conclusion

We have created a new expressive musical instrument that maps movement to sound. The prototype follows the post-digital approach inspired by reacTable synthesizer [2]. It combines the physical and tangible with the digital. In this way, we achieved an instrument that is intuitive to use and gives enough challenge to master. By the nature of the instrument, it can be modified and adjusted to fit any musical genre. The MIDI interface makes it easy for musicians to integrate it without needing to code.

## 9 Funding and Competing interests

This research has been supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS22/172/OHK3/3T/13 and by Technology Agency of the Czech Republic, project No. CZ.02.1.01/0.0/0.0/16 019/0000765 (Research Center of Informatics).

## References

1. Bovermann, T., Groten, J., deCampo, A., Eckel, G.: Juggling sounds. In: Proceedings of the 2nd International Workshop on Interactive Sonification (2007)
2. Jordà, S., Geiger, G., Alonso, M., Kaltenbrunner, M.: 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. pp. 139–146 (2007)
3. Leischner, V., Míkovec, Z.: Spatial audio music player for web. In: RoCHI. pp. 125–129 (2021)
4. Leischner, V.: Vr 360 with custom made controller in java (Nov 2022), `https://www.youtube.com/watch?v=C-CQz4nFYtA`

# Investigating the Role of Usability, User Experience, and Aesthetics for Industrial Human-Machine Interfaces

Annika Hoffmann, Elze Tomasiunaite

Department of Visual Engineering, University of Applied Sciences Dresden, Friedrich-List-Platz 1, Dresden, 01069, Saxony, Germany

**Abstract.** This paper investigates whether an aesthetic design can improve the user experience and usability of industrial human-machine interfaces (HMIs) used in factories, given the current demographic changes and the shortage of skilled workers. A simulation study was carried out with a HMI of a simulation of a production facility. After an evaluation, the HMI was redesigned and a second study was conducted with the same subjects. The method of doing an online simulation study instead of carrying it out at the actual production plant was chosen due to safety reasons. The results showed that the HMI's new design was more user-friendly as both UEQ and SUS scores were higher. The UEQ scores are used to make statements about the user experience and the SUS score is used to make statements about the usability of the affected HMI. In the second study the UEQ scores were higher in every category and the SUS score could also be raised from 45,7 (average) to 76,4 (good).

## 1 Introduction

Developments such as demographic change and digitization are intensifying the shortage of skilled workers at companies in industrial manufacturing. To compensate this and ensure a versatile production, an easier induction for those workers is beneficial. One possibility to realize this is to investigate the accessibility and learnability of industrial HMIs by which operators can monitor and control plant and machinery. Usability, user experience, and aesthetics play a major role here. The correlation of usability and aesthetics is explained in an article written by Lima and Wangenheim [7]. To further investigate this subject, research on the question "Can usability and user experience when interacting with an HMI in an industrial production environment be improved through aesthetic design?" was done. It is possible to take measurements to answer that question using standardized questionnaires. Subsequently, a revision of the HMI design can be made on the basis of insights gained during the evaluation of the questionnaires and especially noted comments of the test candidates due to keeping a think-aloud protocol and free text fields. In this case, there were user tests carried out on the HMI and questionnaires given to the subject group to evaluate their impression and perception about the HMIs look, usability and the experience using it.

## 2 Related Work

The ISO 9241-11 standard [12] explains the concept of usability and its application when using interactive systems, products and services. ISO 9241-110 [11] describes interaction principles between a user and a system and the application of them and general design recommendations. ISO 9241-112 [13] establishes ergonomic design principles for interactive systems regarding the perception and understanding of presented information. Another related paper is a case study [4] in which the tool eXperience Capturer is used to evaluate the interaction between employees and industrial HMIs was used. The study shows that using the eXperience Caputurer facilitates creating new designs that improve the user experience. Also the paper "Human-Machine Interface: Design Principles of Visual Information in Human Machine Interface Design" [8] gives some principles which should be noted in visualization design and which provide the visualization design of HMIs a fundament. Still, there is a research gap concerning targeted investigations of usability and UX for industrial interfaces. In order to receive comparable results the present study resorted to using two questionnaires, the first one being the "System Usability Scale" (SUS) [6] to measure the usability of the HMI. The second questionnaire is the "User Experience Questionnaire" (UEQ) [2]. Which is used to measure the user experience.

## 3 Studies Regarding the HMI and it's New Design Implementation

In the following, the methodology, implementation, and results of the user studies before and after the HMI redesign are described in detail. The procedure consists of three steps: First the initial design is being evaluated through a simulation, then a redesign of the HMI is being executed and lastly a second evaluation takes place in the same format, only with the new design. For the first study 19 subjects with no prior knowledge and 4 subjects with prior knowledge were selected. The identical group participated also in the second study to evaluate the new design and reassure that their suggestions and critique were interpreted correctly. The usage of a completely new subject group would be planned for a further step in this research. Both studies were conducted outside the facility, on an identical simulation-twin on a personal computer, because the system is susceptible to faults and malfunctions and thus could be easily damaged. The simulation was implemented with the application "NX - Mechatronics Concept Designer" and represented both - the entire station and its entire functionality. To provide full control of the simulated station, the HMI was coupled with Siemens "TIAV16" application. The HMI for the simulated station was operated using the mouse instead of direct touch, which could lead to minor deviations in the later evaluation.

### 3.1 First Study

In the following the process of conducting the first study to determine the usability and UX operating the initial state of the HMI are presented.

**Original design evaluation** Subjects were accompanied by a supervisor and lead through the following steps: 1. A brief introduction to the facility and the specific system with the HMI was given. 2. Three different use cases were presented to the subjects without showing the required controls on the HMI interface: a) activate the automatic mode; b) carry out the steps of the automatic mode manually; c) empty the magazine regardless the mode. 3. Subjects were left to conduct every use case on the simulation without any further help or influence, but were asked to think aloud and give live feedback. The supervisor kept a protocol of the participants' comments and tracked time for each task. 4. After the successful completion of all three use cases, the UEQ and SUS questionnaires were handed out.

**Results of the Study on Original HMI Design** Most criticism was shown on the general interface, for example "it's outdated and in need of development", "cluttered". Subjects stated that the colour scheme was confusing and proportions of the elements distracting. There were few to no explanations and some terms were misleading. Still, the HMI's usability and UX was perceived as being sufficient and the general handling improved after the participants had become familiar with the task and interface. Subjects also left suggestions for improvement. This mainly concerned the structure of the menu, the general design in terms of colours, button sizes and layout, as well as designations, general usability, and user experience. The HMI reached a SUS score of 45,7 (less than the average of 68), which confirms bad usability [5]. When compared to the benchmark test provided by the UEQ [2], the HMI was far below average in every category (see Fig. 1; lower line).

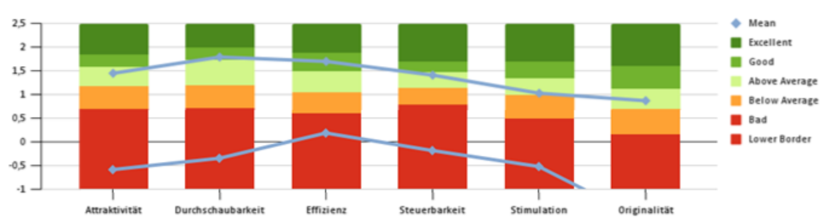



**Fig. 1.** Lower line shows UEQ benchmark score regarding the initial HMI and upper line shows the score regarding the new HMI [2]

### 3.2 Re-Design of the Initial HMI

On the basis of the insights retrieved from the first user study with the initial HMI design, the new design concepts were prepared. The process of designing took place on the browser-based interface designing tool Figma [1], which allowed for easy collaborative work and didn't charge fees. The interaction implementation took place on AdobeXD [3], which provides a link to be shared with the subject group to conduct the second testing. At first, main views were redesigned. Second, due to the probands requests, windows help and login were newly created and thus added to the second user study phase. The factory employees set following requirements for the new design: all functionalities must continue to be represented, as well as the manufacturer's labelling and other copyright elements. Throughout the designing process guide rules were established based on the ISO norms, papers mentioned in related works and user feedback: specific colour coding (active buttons in blue, running text in grey, footer information in light grey, error in red, warning in yellow, correct in green [10]), fixed font style and size, clean line-up of elements (like a row of buttons). These rules help prevent distraction [13] by providing a clean layout, plenty of whitespace, tidiness and consistency [13], which were missing on the initial HMI design according to the first user study. The new design solved complaints about the unstructured main menu by focusing on compactness [13]: Menus and some buttons have been merged together, the order of some buttons was changed for logical purpose (see Fig. 2 & 3). Texts were substituted by new symbols on the controls to support faster reaction and navigation. All in all, the goal was to emphasize orderly and clear design, which are main notions in classical aesthetics [9].

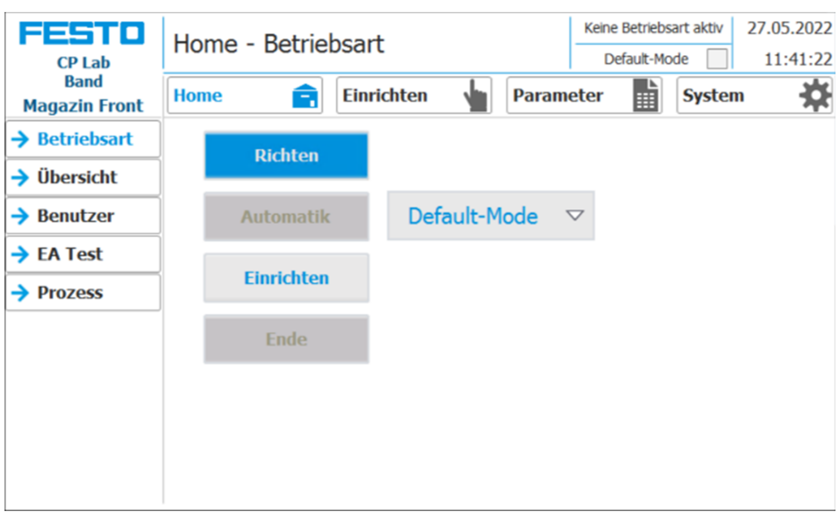



**Fig. 2.** Initial HMI startscreen

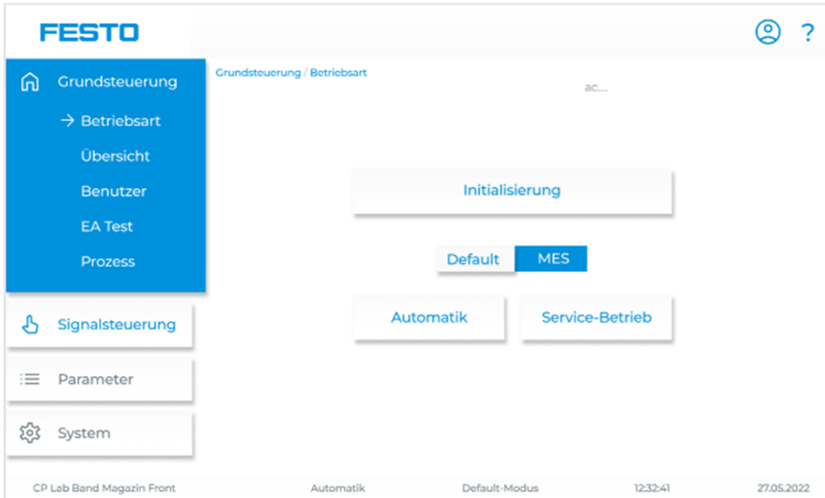



**Fig. 3.** New HMI startscreen

### 3.3 Second Study

The second user study was conducted to reveal whether the changes in design improve the usability and user experience or not.

**Re-Design Evaluation** The second study was performed nearly identical to the first one, containing only slight differences: The evaluation object was the new design, the probands had already prior knowledge about the system and the interface from the first study (to assess the design changes and to speed up the testing phase), the new design prototype was built on AdobeXD and not connected to the previously mentioned simulation program. Hence, the aftermath to every interaction with the machine could not be observed (to speed up the testing and concentrate only on the design).

**Results of the Second Study** The feedback appeared diverse, still most of it was affirmative: "the new design appears to be tidy and lucid", "appealing and modern", "the help window came in handy", "the menu structure has improved", to name a few. On the other hand, the usage of confusing terms and the hidden help window were prominently named as negative aspects. The prototype of the new HMI reached a SUS score of 76,4. This result could be interpreted as good [5]. Across all categories of the UEQ with scores now ranging from 0.868 up to 1,789 the HMI could be considered good compared to the neutral scores from the previous measurement [2]. The UEQ benchmark test showed all scores being above average with some of them even considered good (see Fig. 1; upper line).

## 4 Discussion

The questionnaires show that the original HMI design has major flaws both in usability and user experience. The recorded time in which the users solved the use cases in the second study resulted in being shorter than the time recorded in the first round. However, this statement has to be viewed critically due to the users' knowledge about the routine in solving those cases brought from the study with the initial HMI. The scores SUS and UEQ measured clear improvement in both usability as well as user experience after changing the HMI design (see Fig. 4). In conclusion, the studies proved that user experience and usability of an interface for human-machine-interaction can be improved through layout and content redesign based on studies and research. These new insights can serve as the basis for further work on the design and interactions.


**Fig. 4.** Comparison of the SUS Scores: the left turquoise bubble shows the initial HMI's score, the right the redesigned HMI's [14]

## 5 Conclusion and Outlook

This paper explored improving industrial HMIs' usability and user experience through more aesthetic design. Two simulation-based studies compare the current HMI design with a potential new design, using questionnaires to assess System Usability Scale (SUS) and User Experience Questionnaire (UEQ) scores. Results show that redesigning with good usability, user experience, and feedback principles leads to higher SUS and UEQ scores, improving accessibility and learnability, which would facilitate a versatile production, especially given the demographic changes and shortage of skilled workers in the industry. Further studies should use the same simulation software for better comprehensibility about the facility to gather feedback from new test subjects without prior knowledge, refining the new design based on their feedback.

**Acknowledgements** This work has been supported by the European Regional Development Fund and the Free State of Saxony (project no. 100602780).

## References

1. Figma: the collaborative interface design tool., https://www.figma.com/
2. User Experience Questionnaire (UEQ), https://www.ueq-online.org/
3. Was ist Creative Cloud für Unternehmen — Adobe Creative Cloud für Unternehmen, https://www.adobe.com/de/creativecloud/business/enterprise.html
4. Aranburu, E., Erle, G.L., Gerrikagoitia, J.K., Mazmela, M.: Case Study of the Experience Capturer Evaluation Tool in the Design Process of an Industrial HMI. Sustainability 12(15), 6228 (8 2020), https://www.mdpi.com/2071-1050/12/15/6228/pdf?version=1596439688
5. Bangor, A., Kortum, P., Miller, J.H.: Determining what individual SUS scores mean: adding an adjective rating scale. Journal of Usability Studies archive 4(3), 114–123 (5 2009), http://uxpajournal.org/wp-content/uploads/pdf/JUS_Bangor_May2009.pdf
6. Brooke, J.H.: SUS: A 'Quick and Dirty' Usability Scale. CRC Press eBooks pp. 207–212 (6 1996)
7. De Souza Lima, A.L., Von Wangenheim, C.G.: Assessing the Visual Esthetics of User Interfaces: A Ten-Year Systematic Mapping. International Journal of Human-computer Interaction 38(2), 144–164 (6 2021)
8. Gong, C.: Human-Machine Interface: Design Principles of Visual Information in Human-Machine Interface Design. International Conference on Intelligent Human-Machine Systems and Cybernetics (8 2009)
9. Hartmann, J., De Angeli, A., Sutcliffe, A.: Framing the user experience: Information biases on website quality judgement. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. p. 855–864. CHI '08, Association for Computing Machinery, New York, NY, USA (2008), https://doi.org/10.1145/1357054.1357190
10. Hossain, A., Zaman, T.: HMI Design: An Analysis of a Good Display for Seamless Integration between User Understanding and Automatic Controls. 2012 ASEE Annual Conference  Exposition (9 2020), https://peer.asee.org/21454.pdf
11. ISO9241-110:2020: Ergonomics of human-system interaction — Part 110: Interaction principles. Standard, International Organization for Standardization, Geneva, CH (May 2020)
12. ISO9241-11:2018: Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts. Standard, International Organization for Standardization, Geneva, CH (Mar 2018)
13. ISO9241-112:2017: Ergonomics of human-system interaction — Part 112: Principles for the presentation of information. Standard, International Organization for Standardization, Geneva, CH (Mar 2017)
14. Smyk, A.: The System Usability Scale  How it's Used in UX - Thinking Design (12 2021), https://medium.com/thinking-design/the-system-usability-scale-how-its-used-in-ux-b823045270b7

# Using optically illusive architecture to navigate users in Virtual Reality

Vojtěch Radakulan

Faculty of Electrical Engineering, Czech Technical University - Department of Computer Graphics and Interaction, Czech Republic

**Abstract.** There are several aspects of physical architecture which are used to communicate to visitors where they should go and where not, without using graphical signs. Some of these elements, like geometrical properties of hallways or entrances could be also used in virtual reality, but with several differences related to the nature of VR and its differences from physical architecture. The idea is to look into a methodology of current solutions in physical architecture and discuss possible usage and adjustments in VR.

## 1 Introduction

There are many optical illusive adjustments used in art and architecture to induce different effect of such object in order to appear differently. Such optical refinements were used since antiquity to either make art and architecture look more immersive, or to emphasise certain aspects of the physical object. The most prominent example is a deliberate curvature of Parthenon in [**?**] sir as in figure 1, which was optically refined to look more straight. [1] This was later used after broader usage of perspective to make illusive spaces where physical building was not possible, as for example in Palazzo Spada by Donato Bramante. [2] This was later described as a forced perspective, which is a technique which uses optical illusions to induce an impression that some object is in a different position, scale or rotation than it actually is. Optically illusive refinements induces viewers spatial perception to misrepresent the reality. When using such illusions, we can lure the viewer to change the vantage point in order to understand the illusion or know the true state of the object. This behaviour is thus governing peoples position in space.[4] When are users exposed to such an illusion, they tend to spend more time in the area from which the illusion takes place.[4], which can be used to manipulate viewers location in VR. The question is however how to be sure to know that the illusion itself is the key factor to change of viewer position. One such indicator could be a use of EEG and more particularly N400 phenomenon first used in linguistics [5], later in 2D graphics and art [6] and most recently in VR environment. [8]. Other possibility how to measure such behaviour is to use psychophysical methods from the area of spatial cognition. [3]

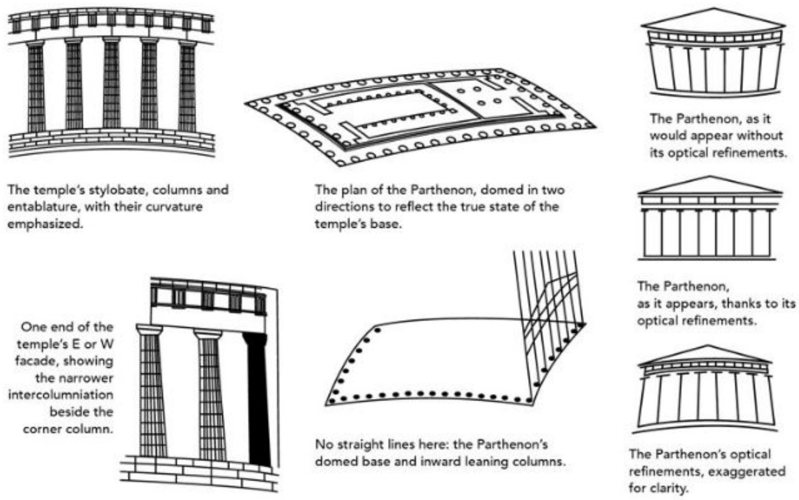



**Fig. 1.** Graphical exaggeration of optical refinements on Parthenon

## 2 Methods of behaviour evaluation within VR

EEG evaluation of such behaviour is great for first encounter of a viewer with such an illusion. However when encountered repetitively, the surprise from deciphering the illusion is less apparent till nonexistent, because viewer will adapt to knowing the illusion. Other method is to use psychophysical methods to have more precise answer from more trials. From our previous work we also made a qualitative research towards understanding spatial illusions in architecture. In our study, our aim was to determine how the properties of a distorted hallway (e.g., angle of the walls, amount of distortion) influence the perceived illusory space. We created a simplified VR model of an illusive hallway , as on figure 2 inspired by a historical example designed by Italian architect Francesco Borromini, and conducted a qualitative pilot study. The interviews revealed that virtual movement (especially changing the user's vantage points towards the sides) and hallway geometry (the angle of hallway walls) were considered to be the most important factors for the illusion.[7]

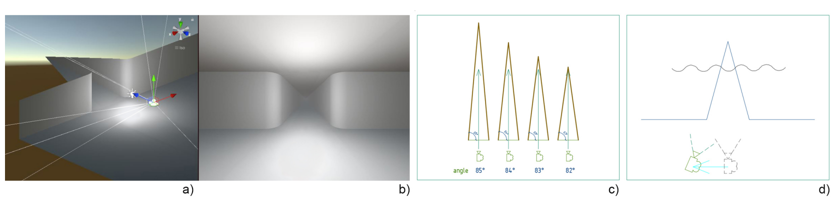



**Fig. 2.** 2a,b) Simplified VR representation of an illusive hallway (orthographic and perspective view) 2c) Example of different angles between perceived and actual walls of the hallway 2d) Virtual movement of the participants and distance fog (wave line), Source: Personal archive

## 3 Pilot study investigating participants detection techniques.

Participants wearing a head-mounted display (HMD) experienced five distorted virtual hallways using the forced perspective technique, and five regular corridors. They were placed in the central point in front of a hallway, where the illusion always works and were asked to use a joystick to virtually move to the left side and signal when they detect whether they are seeing actual or distorted corridor. As in figure 3, one can compare two behaviours of same participant. In one case, when presented actual corridor, participant moved just about one meter towards left and then determined that was an actual corridor. However, when presented distorted corridor, user first spent about 4 seconds observing it, and later moved back and forth in order to determine whether that is an illusive corridor. This behaviour however was usually presented only during first few trials and one explanation for it is the paralax phenomenon. The other takeaway from piloting was the fact that first few participants were tested with heavily reduced experience (there was light only from the environment which consisted from gradient of three colours). When presented such experience, users were unable to determine anything and failure to identify the corridor was almost exactly random (49,3 percent). We had to introduce some of refinements, especially light and steps to cover the floor transition from flat to steep distorted slope, as on image 4 . After that, the fail rate dropped to 7,2 percent.

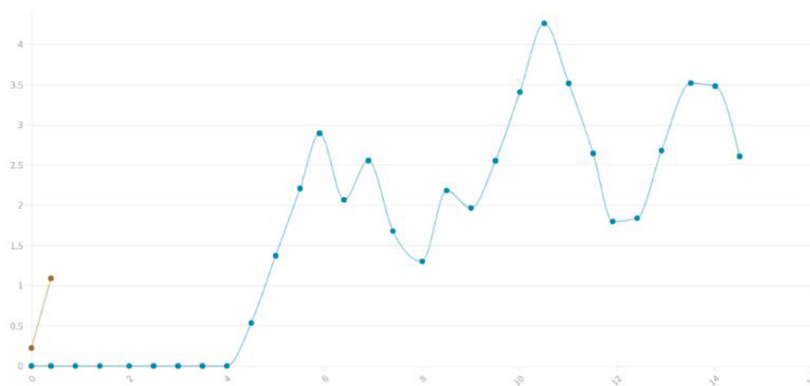


**Fig. 3.** X: time, Y: distance traveled from origin. Blue: Fake Corridor, Brown: Normal Corridor

## 4 Future work

In our ongoing work, we are currently designing an experiment based on psychophysical methods to determine how the factors of hallway geometry influence views behaviour. We would like to focus on angle of walls as identified in former pilot study. Subsequently we would also like to find out to what extent the size of the area in which the illusion can be experienced by the visitor changes based on hallway geometry factors.

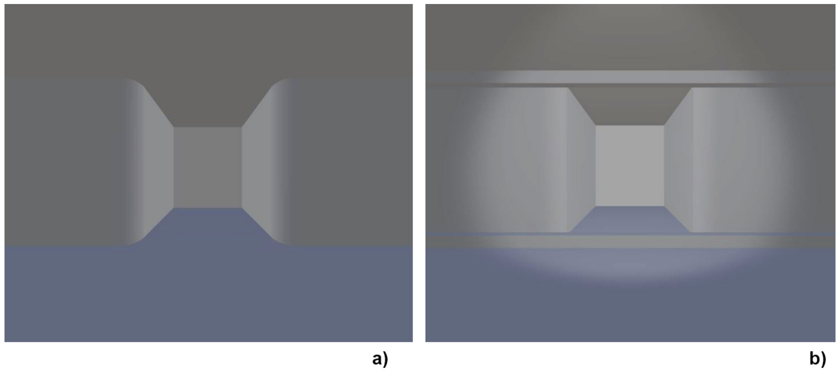


**Fig. 4.** A: maximally reduced corridor B: Adjustment of reduction


## Acknowledgements

The research has been supported by the project Research of new interaction methods in special environments funded by grant no. SGS22/172/OHK3/3T/13 (FIS 161 - 1612233C000). This research has been supported by the TACR project the project Research Center for Informatics (reg. no. CZ.02.1.01/0.0/0.0/16 019/0000765) supported by EU.


## References


1. Coulton, J.J.: Ancient Greek architects at work: Problems of structure and design. Cornell University Press (1982)
2. Dunning, W.V.: Changing images of pictorial space: a history of spatial illusion in painting. Syracuse University Press (1991)
3. Frankenstein, J., Kessler, F., Rothkopf, C.: Applying psychophysics to applied spatial cognition research. In: German Conference on Spatial Cognition. pp. 196–216. Springer (2020)
4. Hosseini, S.V., Alim, U.R., Oehlberg, L., Taron, J.M.: Optically illusive architecture (oia): Introduction and evaluation using virtual reality. International Journal of Architectural Computing p. 14780771211016600 (2021)
5. Kutas, M., Federmeier, K.D.: Thirty years and counting: Finding meaning in the n400 component of the event related brain potential (erp). Annual review of psychology 62, 621 (2011)
6. Markey, P.S., Jakesch, M., Leder, H.: Art looks different–semantic and syntactic processing of paintings and associated neurophysiological brain responses. Brain and Cognition 134, 58–66 (2019)
7. Radakulan, V. Frankenstein, J.S.D.M.Z.R.K.: Towards understanding spatial illusions in architecture - a pilot study exploring factors influencing illusive perspectives (2022)
8. Tromp, J., Peeters, D., Hagoort, P., Meyer, A.S.: Combining eeg and virtual reality: The n400 in a virtual environment. In: the 4th edition of the Donders Discussions (DD, 2015) (2015)

# Speed and Required Precision of Grabbing Physical Spheres in VR

Uršuľa Žákovská

Faculty of Electrical Engineering,
Czech Technical University in Prague, Czech Republic

**Abstract.** Simultaneous interaction with a physical object and its virtual twin, whether in Virtual reality (VR) or Augmented reality (AR), usually requires high precision and likeness of the two representations. However, in the case of perishable and not completely solid physical objects such as food, it is necessary to expect some inaccuracy in shape, size or even position. This paper presents multiple approaches for more relaxed requirements which are still acceptable to the user when interacting with the food. The speed of the interaction is not essential for this use case. Still, it serves as an easy-to-understand baseline comparison of how natural it feels for the user depending on what he interacts with, real, virtual or both representations of the object simultaneously.

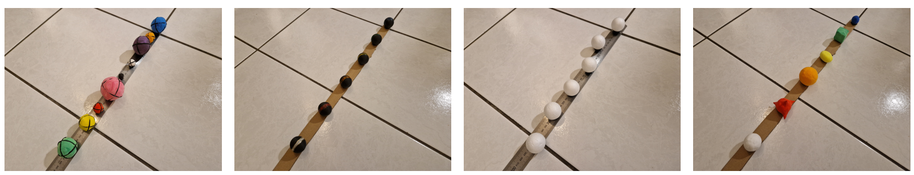
## 1 Introduction

For VR to be an effective tool for gaining a new practical interaction experience, it must be able to replicate the natural physical interactions that occur in the real world. For example, in VR-based training simulations, users need to manipulate virtual objects naturally and intuitively, as this will enhance their ability to learn and retain information. High precision and likeness of the virtual and physical representations of an object in VR are essential for achieving a believable and immersive experience. It allows the user to clearly understand the object's properties, perform actions, make predictions, and have accurate and responsive feedback in the virtual environment. It is worth noting that while visual modalities can be replicated well in VR, it is more challenging to replicate the texture, taste, or smell of objects via touch or olfactory modalities. This is why incorporating physical objects in Extended reality (XR) experiences can be beneficial or even essential, depending on the purpose of the experience.

However, when it comes to food, standard ways of tracking these objects based on their physical appearance can prove too hard of a task, as shape, colour, size, and other properties are often very diverse, even across one ingredient. Let us assume we managed to do it by either oversimplifying and scrutinizing the physical appearance, creating a massively robust tracking system, or we do not track the food but the human. This experiment implements the last option. It requires predicting and assuming what happened to the food based on observed human interaction and tendencies with the food. We must persuade human participants that eating this food is safe and desirable. This paper explores the limits of how lenient we can be when replicating the food/object's properties and still be able to retain immersion and believability.

## 2 Methodology

Before the experiment, we decided to focus on two questions. First, what is the ideal size of such an object/food or its part? Second, what can be changed from reality, and to what extent? When it comes to the second question, we will explore size, shape, offset and weight change concerning the physical sphere. Does it bother or faze the player if the sphere is smaller, bigger, flatter, squarish, heavier, lighter, or offset? Wherever possible, the spheres were made from a squishy material to reflect the consistency of real food (Figure 1a, 1c, and 1d).



(a) Sizes (b) Weights (c) Offset (d) Shapes

Fig. 1: Experiment Tracks: (a) objective ideal size for interaction (randomized); (b) subjective ideal weight (in order); (c) objective maximum offset (randomized); (d) subjective reaction to morphed shapes (randomized)

During the development, we found out that the standard physical-based hand-tracking interaction toolkit would not work well with physical objects because the object (sphere) would cover one or more fingers and consequently cause the dropping of the virtual sphere. This led to the implementation of simplified grabbing interaction where once the sphere is grabbed, it stays attached to the thumb. This solution came from observing human-food interaction where people do not let go of the food once they pick it up and also they try to minimize the area of contact.

(a) two (b) one (c) sometimes two

Fig. 2: The number of fingers touching the sphere according to the standard hand-tracking interaction toolkit. Too unstable for our purposes.

Part of the preliminary research was the maximum bite size (Figure 3a). Participants (10 of them) were asked what is the largest sphere they would naturally eat in one bite. Based on that, we chose the sphere size for Weight, Offset and Shape tracks. Afterwards, when all the props were prepared, the main experiment with a different group of participants was conducted. Participants (8 of them) were first asked to answer a set of questions about their VR experience and handedness to ensure that only right-handed and VR experienced participants participate to reduce the variables. Each participant received a VR headset (Oculus Quest) and was seated in front of a table with bowls (Figure 3b). The virtual scene was aligned to the real (physical) one through the passthrough mode prior to this.

(a) different sizes in order

(b) experiment setup

Fig. 3: First two steps of the experiment: (a) questionnaire and maximum bite size selection (b) VR headset equipping

As mentioned above, the experiment was divided into four tracks (Size, Weight, Offset and Shape). However, the first stage (Size) was further separated into three parts depending on what spheres participants were interacting with:

(A) virtual spheres
(B) real and virtual spheres simultaneously
(C) only real (physical) spheres

That means, overall, there were six different tracks:

1. Size A
2. Size B
3. Size C
4. Weight
5. Offset
6. Shape

We used three lion heads instead of one for the Weight and Shape tracks. These were the tracks about exploration, not speed. We gave participants a primary goal: correctly categorize given word by feeding one of the lionesses (Figure 4c).

While testing the setup, people did not respond to regular bowls well (Figure 4a). More specifically, because they were not interacting with the actual food but only with props, their behaviour did not follow the same behavioural patterns as was previously observed in human-food interaction (mainly grabbing spheres differently, e.g. into a fist). For that reason, the neutral white bowls were swapped with lion heads (Figure 4b) to enforce the feeling of feeding someone.


(a) old version: 3 bowls (b) new version: 1 bowl (c) new version: 3 bowls

Fig. 4: Activities during the experiment: (a) and (c) show exploration tracks of categorizing given word while (b) shows a track focused on speed when the sphere had to be inserted into the lion head as fast as possible.

## 3 Results & Conclusion

In the preliminary part, we found out that the largest sphere (eaten in one bite) among the nine we offered (15-55mm in $\varnothing$) is a sphere of 35mm in diameter (yellow sphere in Figure 3a). Consequently, spheres for Weight, Offset and Shape tracks were prepared in this size. The first three tracks (Size A, B and C) show that participants are faster the bigger the sphere is regardless of whether we are measuring real (physical), virtual or both spheres simultaneously. Also, they were roughly twice as fast when interacting only with real spheres in comparison to simultaneous real and virtual spheres (Table 1 and Figure 5).

| Size $\varnothing$(mm) | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 |
|---|---|---|---|---|---|---|---|---|---|
| Real | 0.38 | 0.32 | 0.31 | 0.31 | 0.29 | 0.31 | 0.30 | 0.28 | 0.25 |
| Virtual | 2.00 | 1.86 | 1.51 | 1.32 | 1.05 | 1.04 | 1.01 | 0.98 | 0.89 |
| Both | 3.72 | 3.17 | 3.15 | 3.05 | 3.03 | 2.58 | 1.97 | 1.86 | 1.64 |

Table 1: Measured time in seconds: Average time it takes to move a sphere from initial touch to placement into the lion's mouth

For the Weight track, spheres were prepared in ascending order from the lightest to the heaviest (8g, 17g, 25g, 34g, 43g and 52g). Some participants were so focused on the separation activity that they did not notice the weight difference. When that happened, they were asked to do the Weight track again, this time aware of their weight variance. The majority said that 25g is the most appropriate weight for food of this size, and 34g was too heavy to be a food. Just for reference, a sphere of the same size made from date paste which is one of the heavies edible ingredients weighs 28g. Another interesting thing was that participants thought the sphere was heavier in VR compared to holding the same sphere in real life.

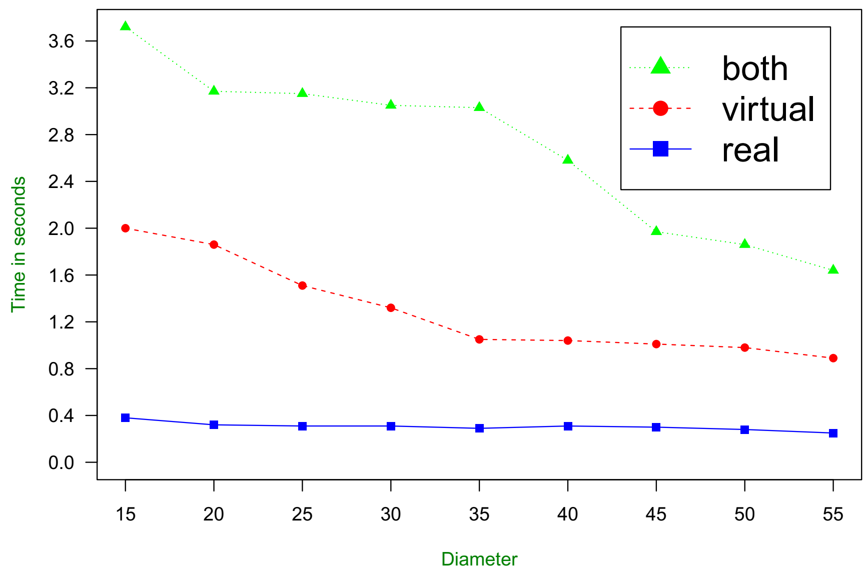



Fig. 5: Graph visualization: Average time it takes to move a sphere from initial touch to placement into the lion's mouth

Offset track with 4 spheres of 0, 5 or 10mm offset had so many invalid runs with each participant that it was impossible to conclude anything. The biggest hurdle was aligning the real and virtual environments with enough precision. The participants could grab the spheres even when the offset exceeded 25mm, but it required some contact-based blind grabbing. The six selected morphed shapes included default, smaller, bigger, squarish, spikey and squeezed (flatter) spheres. Participants noticed a difference for all morphs except smaller and spikey spheres, but they all enjoyed the experience of exploring the shapes, and it did not get in the way of interacting with them.

## Acknowledgement

The research has been supported by the project Research of Modern Computer Graphics Methods 2022-2024 funded by grant no. SGS22/173/OHK3/3T/13 (FIS 161 – 1612234C000).

# ReFlex - A Framework for Research on Elastic Displays

Mathias Müller & Dietrich Kammer

Department of Informatics / Mathematics, University of Applied Sciences, Dresden Germany

**Abstract.** **Elastic Displays** as a specific type of **Shape-Changing Interfaces** provide haptic feedback by using the deformation of the screen for interaction. There have been several prototypes showcasing different use cases, visualization techniques and interaction metaphors. In this paper, we present an overview over our efforts to create a modular and extendable software framework to support research in the context of this specific type of interface. The main goal is to provide a common and easy starting point to streamline research on this technology. The **ReFlex framework** is based on a **service-oriented client-server architecture** and uses modern **web technologies** for platform-independent development. Furthermore, it provides tools for **configuration** and **calibration** to simplify the hardware setup. **Development tools** as well as plugins and templates for common client technologies provide a starting point for rapid prototyping concepts for Elastic Displays. Finally, we provide insights on our experiences with framework as part of student projects and describe our ideas for future research on possible framework extensions and technological enhancements in the context of Elastic Displays.

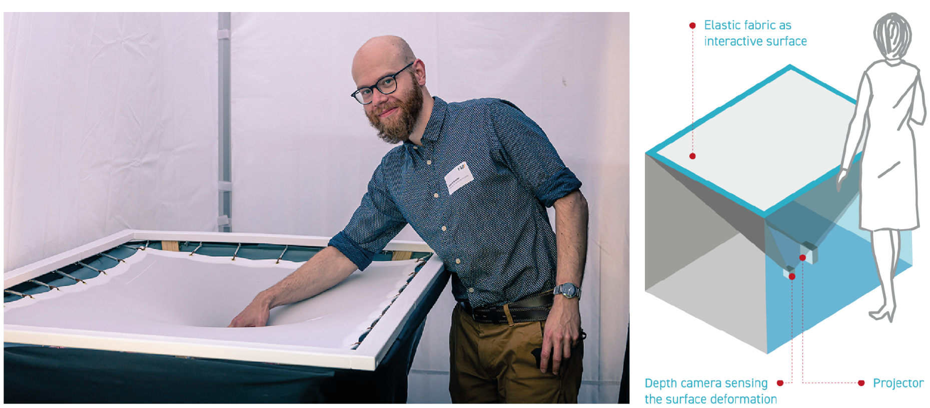



**Fig. 1.** *Left*: Prototype of the **Elastic Display** as interactive table-top, *Right*: Schematic hardware setup.

## 1 Introduction & Related Work

Haptics in human-computer interaction is an important field of research. Current touch interfaces do only insufficiently use the potential of the human hands in regard of expressiveness and capabilities. Instead of purely visual communication, **Shape-Changing Interfaces** enable rich haptic experiences and explore novel ways to interact with virtual content [1]. **Elastic Displays** are a specific type of **Shape-Changing Interfaces** that focus on the deformation of an elastic surface for interaction. Usually, an elastic fabric is used as projected screen, which can be pushed or pulled at different locations (cf. Fig. 1). The resistance from the fabric is a natural haptic indicator of the interaction depth, providing passive haptic feedback. In order to reduce the mechanical complexity of the hardware, there is no active deformation of the surface, opposed to **Actuated Displays** [8]. The *Khronos Projector* prototype was one of the first concepts demonstrating the haptic exploration of an information space by interactively deforming the elastic surface [2]. Since then, there has been a magnitude of different prototypes showcasing different concepts of **Elastic Displays**. This includes examples for manipulating spatial content with the *Deformable Workspace* [18], navigating **graph visualizations** and computer tomography imagery [20], using **physical interaction metaphors** [11, 12], exploring **layer-based data** [13], and interacting with complex data in **zoomable user interfaces** [7, 12]. Although most publications refer to prototypes exploring different interaction metaphors and visualization techniques, there has also been some research on the fundamentals of interaction with **Elastic Displays**: *Troiano et al.* conducted a user study to explore how users naturally interact with Elastic Displays and which **gestures** they use [17]. More recently there also has been a study on the capabilities of users in regard to the **sensitivity of depth interaction** when exploring layer-based content [14]. As part of model-building efforts, recent works identify common data types suitable for **Elastic Displays** [4] and develop a **task taxonomy** based on analysis of several prototypes [8].

## 2 ReFlex Framework

To address the fragmented field of research, we wanted to create a software framework which serves as a common starting point to explore the opportunities of this new technology without dealing with low-level issues like querying, processing and interpreting sensor data, or calibration of the system. Therefore, we implemented a **modular framework** using a **client-server architecture** (cf. Fig. 2): the server retrieves data from the depth sensor using a **sensor abstraction layer** which implements a number of common depth cameras and can be easily extended to be used with different sensors. The tracking server also features **filtering** methods to cleanup and smooth the incoming data, data **processing** algorithms to compute **interactions** on the surface, and methods for sensor-display **calibration**. Clients receive touch event data via **WebSockets** using a custom protocol or an implementation of the **TUIO protocol** [6]. The server also features a web-based user interface where current the current calibration, sensor data (cf. Fig. 2) and interactions (cf. Fig. 3) are visualized. The framework is implemented using **.NET** [10], so it can be used on all current desktop systems, and the tools can be packaged as **Electron** [16] desktop applications. The framework also provides several tools to facilitate the development of applications, such as an **emulator** (cf. Fig. 3) and tools for **recording** and **replaying** sensor streams, including sequences from the emulator. Furthermore, the framework offers templates for client applications using **Angular** [5], **React** [9], **Vue.js** [19] for web-applications or WPF/.NET Core or an **Unreal Engine 4** [3] plugin for desktop applications.

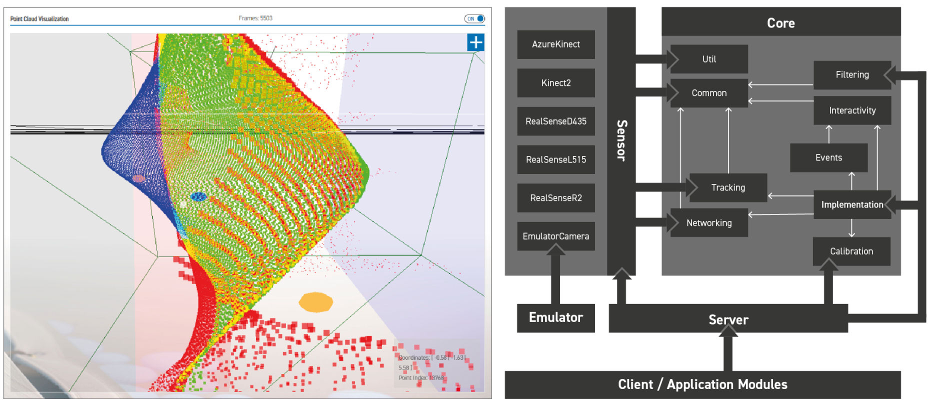



**Fig. 2.** *Left*: Visualization of the sensor data as point cloud in the user interface of the **ReFlex** tracking server component. *Right*: Architecture diagram of the **ReFlex framework**. It uses a distributed **client-server-architecture** and offers a common interface for different sensors and separate modules for processing sensor data, calibration, computing interactions, and broadcasting data.

## 3 Lessons Learned

The framework has been used in **several student projects** so far. Student feedback provided valuable insights into different aspects for refinement and optimization, as well as compatibility with different hardware setups. It also served as proof-of-concept for the core idea to facilitate the development of applications for this novel technology. The provided developer tools, especially the emulator and application templates, have helped to **rapidly explore use cases** and implement the basic functionality without the need to constantly testing on the **Elastic Display** itself. The students implemented gestures and basic interactions in their **familiar development environment**, and only added final optimizations, parametrization and bug fixes for special cases at the end when testing with the final hardware setup. Display setup can also be significantly simplified with the tools provided by the framework, especially **loading and saving different configurations** allows for rapidly switching between different prototypes. Within the different iterations of the framework over the last years, **extendability and modularity** have been identified as key features. Especially with the development of **different tracking and processing approaches**, as well as integrating different sensors, it was possible to proof that this goal has been achieved to a large degree.


**Fig. 3.** *Left*: Visualization of the interactions detected by the framework using the emulator as input source. The interactions are labeled with a *touch id*, and have attributes like *position (including depth)*, *confidence*, and *trace*. *Right*: User Interface of the **ReFlex emulator** tool with two emulated deformations: push (*grey*) and pull (*yellow*). The diameter of the circle depicts the depth. The client web app is displayed in the background of the emulator, visualizing the calculated extrema broadcasted by the server.

## 4 Future Directions

One of the major constraints of **Elastic Displays** is that haptic feedback is limited to the passive force resulting from the deformation of the fabric with no simple option for programmatically generating haptic effects to enhance user interaction. However, concepts for providing active haptic feedback using vibration motors, have been explored recently [15]. One future direction is to conduct an in-depth evaluation of the use of **vibro-tactile feedback** using the **ReFlex framework** for a user study exploring different vibration patterns in a real world scenario with **Elastic Displays**. Furthermore, detailed measurements of the fabric could also be used to create a better understanding on how vibrations are transported over the surface. Another open question is to use methods from materials sciences to measure **surface tension and material properties** of different fabrics to create a standardized model for describing the surface properties and to allow accurate reproduction of the system and its behaviour. For the framework, there are attempts to extend **gRPC** communication for integrating external processing components as **micro services**, e.g. for **integrating**

**machine learning** for better prediction and filtering of touches or alternate image processing techniques for detection of touches.

### Acknowledgements

This work has been supported by the European Regional Development Fund and the Free State of Saxony as part of the ZELASTO project (project no. 100376687).

## References

1. Alexander, J., Roudaut, A., Steimle, J., Hornbæk, K., Bruns Alonso, M., Follmer, S., Merritt, T.: Grand challenges in shape-changing interface research. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. p. 1–14. CHI '18, Association for Computing Machinery, New York, NY, USA (2018), https://doi.org/10.1145/3173574.3173873
2. Cassinelli, A., Ishikawa, M.: Khronos projector. In: ACM SIGGRAPH 2005 Emerging technologies. pp. 10–es (2005)
3. Epic Games, Inc.: The most powerful real-time 3d creation tool - Unreal Engine. https://www.unrealengine.com/ (20014 - 2023), accessed: 2023-03-07
4. Franke, I., Müller, M., Gründer, T., Groh, R.: Flexiwall: Interaction in-between 2d and 3d interfaces. vol. 434 (06 2014)
5. Google: Angular. https://angular.io/ (2010 - 2023), accessed: 2023-03-07
6. Kaltenbrunner, M.: TUIO 2.0 protocol specification. http://www.tuio.org/?tuio20 (2014), accessed: 2023-03-07
7. Kammer, D., Keck, M., Müller, M., Gründer, T., Groh, R.: Exploring big data landscapes with elastic displays. In: Burghardt, M., Wimmer, R., Wolff, C., Womser-Hacker, C. (eds.) Mensch und Computer 2017 - Workshopband. Gesellschaft für Informatik e.V., Regensburg (2017)
8. Kammer, D., Müller, M., Wojdziak, J., Franke, I.S.: New impressions in interaction design: A task taxonomy for elastic displays. i-com 17(3), 247–255 (2018), https://doi.org/10.1515/icom-2018-0021
9. Meta Platforms, Inc.: React - a javascript library for building user interfaces. https://reactjs.org/ (2023), accessed: 2023-03-07
10. Microsoft: .NET — free. cross-platform. open source. https://dotnet.microsoft.com/ (2023), accessed: 2023-03-07
11. Müller, M., Gründer, T., Groh, R.: Data exploration on elastic displays using physical metaphors. In: xCoAx 2015: Proceedings of the Third Conference on Computation, Communication, Aesthetics and X. Universidade do Porto, Porto (2015)
12. Müller, M., Keck, M., Gründer, T., Hube, N., Groh, R.: A zoomable product browser for elastic displays. In: xCoAx 2017: Proceedings of the Fifth Conference on Computation, Communication, Aesthetics and X. Universidade do Lisbon, Portugal (2017), http://2017.xcoax.org/pdf/xcoax2017-muller.pdf
13. Müller, M., Knöfel, A., Gründer, T., Franke, I., Groh, R.: Flexiwall: Exploring layered data with elastic displays. In: Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces. p. 439–442. ITS '14, Association for Computing Machinery, New York, NY, USA (2014), https://doi.org/10.1145/2669485.2669529

14. Müller, M., Stoll, E., Krauss, A.M., Hannss, F., Kammer, D.: Investigating usability and user experience of layer-based interaction with a deformable elastic display. In: Proceedings of the 2022 International Conference on Advanced Visual Interfaces. pp. 1–9 (2022)
15. Müller, M., Kammer, D.: Augmenting elastic displays with active vibrotactile feedback. In: Marky, K., Grünefeld, U., Kosch, T. (eds.) Mensch und Computer 2022 - Workshopband. Gesellschaft für Informatik e.V., Bonn (2022)
16. OpenJS Foundation: Build cross-platform desktop apps with javascript, html, and css — electron. https://www.electronjs.org/ (2021), accessed: 2023-03-07
17. Troiano, G.M., Pedersen, E.W., Hornbæk, K.: User-defined gestures for elastic, deformable displays. In: Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces. p. 1–8. AVI '14, Association for Computing Machinery, New York, NY, USA (2014), https://doi.org/10.1145/2598153.2598184
18. Watanabe, Y., Cassinelli, A., Komuro, T., Ishikawa, M.: The deformable workspace: A membrane between real and virtual space. In: 2008 3rd IEEE International Workshop on Horizontal Interactive Human Computer Systems. pp. 145–152 (2008)
19. You, E.: Vue.js - the progressive javascript framework. https://vuejs.org/ (2014 - 2023), accessed: 2023-03-07
20. Yun, K., Song, J., Youn, K., Cho, S., Bang, H.: Elascreen: Exploring multi-dimensional data using elastic screen. In: CHI '13 Extended Abstracts on Human Factors in Computing Systems. p. 1311–1316. CHI EA '13, Association for Computing Machinery, New York, NY, USA (2013), https://doi.org/10.1145/2468356.2468590

# Digital Reading Stand (DRS)

Kevin Slijepcevic
kevin.slijepcevic@htw-dresden.de

Department of Artificial Intelligence / Cognitive Robotics, University of Applied Sciences, Dresden Germany

**Abstract.** This project shall enable elderly or blind people to be able to conveniently read books, newspapers or other text-sources via magnifiying an image of the text and displaying it on a TV. In the second phase of this project, a text-to-speech-system will read the text to the user. The operation of the DRS shall be as easy as possible, to offer no obstacles to the user. That includes an automatic stream control, automatic shut-down functionality and automatic text-to-speech activation.

## 1 Hardware

This section covers the hardware, which has been used to realize a working prototype. Important criteria for selecting the hardware was the price and availability at the time of development. The DRS consists of two main components: the camera unit and the display/control unit.

### 1.1 Camera unit

The camera unit shall be as small as possible and hand-held. It shall also run a fully functional operating system to execute scripts. Its task is to capture the text with a camera and distribute the resulting video stream via network. A good candidate to comply for these requirements is the *Raspberry Pi Zero 2 W*, running Raspberry Pi OS. The CPU is powerful enough to handle the processing of the video stream, while also being energy-efficient to give the camera unit a good battery life. Also, the small dimensions and cheap price are an adventage, too. To capture the text, a camera is attached to the Raspberry Pi, called *Zero-Cam*. It is a light-weight and cheap camera with a maximal resolution of 1920 x 1080 pixels in video mode. The quality is good enough, though there is room for improvement. In a second iteration, the camera should be replaced with a better model, due to its relevance for capturing the text, which is the main goal of this project.

### 1.2 Powerbank

The camera unit is connected via Micro-USB to a powerbank, which has a capacity of 20.000 mAh and operates as the main power source for the camera unit. The powerbank also acts as the main structure for attaching the camera unit, which is glued into place, as can be seen in Figure 1. In a next evolutionary step of the project, the battery shall be charged wirelessly or with pin-contacts, to make charging as easy as possible for the user.


**Fig. 1.** Fully assembled camera unit with powerbank acting as the main structure and power source

### 1.3 Display/Control unit

The display/control unit is a stationary device, which is connected to a TV via HDMI. Its task is to display the video stream on the TV and to control the stream and camera unit itself. It shall also run a fully functional operating system, which is able to execute scripts. An ideal candidate for this task is the *Raspberry Pi 4 Model B*, running Raspberry Pi OS. The CPU has more processing power, compared to the CPU used in the camera unit, which is needed to display the video stream and run the text-to-speech-pipeline. The communication between the display/control unit and the camera unit is based on TCP/IP and encrypted via a VPN tunnel.

## 2 Processing the video stream

The camera unit captures the video stream with a resolution of 1920 x 1080 pixels and a frame-rate of five frames per second. The stream then gets encoded with the H.264 Codec, encrypted, using *Wireguard*, and then distributed to the network using the *Real-time Streaming Protocol (RTSP)* and *Real-time Transport Protocol (RTP)* with the command line version of the VLC-Player, called *clvc*. The display/control unit then receives the stream, decrypts and decodes it and displays it on the TV. The stream application to view the stream has been realized with a Python script using *OpenCV* and *PyQt*.

The frame-rate of five frames per second has been determined by experimentally using different values and analyzing the resulting video stream and heat dissipation of the CPU. It is a trade-off between energy consumption of the camera unit (the higher the frame-rate, the higher the heat dissipation thus the higher the energy consumption) and the latency and fluidity of the video stream. Due to the main purpose of the DRS, to display static, rarely moving text, a rather low frame-rate has been chosen. This results in good battery life and low heat dissipation of the CPU embedded in the camera unit, while remaining good readability of the captured text.

## 3 Image processing pipeline for Text-To-Speech (TTS)

The text-to-speech pipeline is a important feature of this project. To use the TTS pipeline, the captured text in the video stream has to be transformed into plaintext first, using a *Optical Character Recognition (OCR)* software called *Tesseract*. This generated plaintext then gets inserted into the TTS pipeline, which is still work in progress and will not be covered in this paper. Also a work in progress is the activation of the pipeline. This has been solved manually by processing an image in a script at the time of writing this paper.

### 3.1 Improving OCR output

A major drawback while using a small camera like the ZeroCam is the ability, to output a consistent image quality. The quality highly depends on the lighting conditions and shadow casting. Tesseract has quite substantial problems recognizing text within images, that are unevenly illuminated, as can be seen in Figure 2. To tackle this problem, an image processing pipeline has been implemented, using image manipulation functions in OpenCV: In the first step, the image gets converted to gray scale and dilated, using a kernel-size of 7x7. After that, the image gets blurred with an Median-Blur-Filter. In the next step, the now created and the original image get mixed together, calculating the absolute difference of both. The resulting image then gets normalized and an adaptive threshold is applied. The resulting binary image can be seen in Figure 3, along with the now much improved recognized text by Tesseract.

## 4 Automated stream-control

An important goal of this project are the easy-to-use ergonomics. That means that the user has to interact as less as possible with the devices, while maintaining full functionality. To realize this goal, two scripts have been designed and implemented which are being constantly executed on the display/control unit.

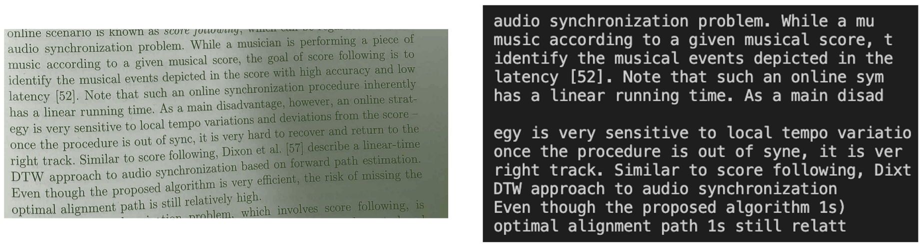



**Fig. 2.** On the left side a snapshot of a book taken with the ZeroCam, on the right side the recognized text by Tesseract

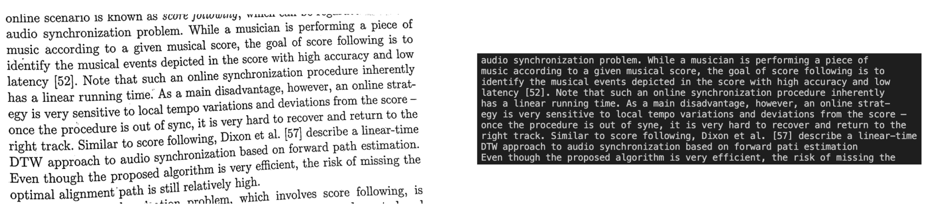



**Fig. 3.** On the left side the same snapshot of the book, processed with the image processing pipeline, on the right side the recognized text by Tesseract of this snapshot

### 4.1 Automated start/stop of the video stream

The stream shall be started and stopped fully autonomous. That means that the user turns on the camera unit and the display/control unit shall pick-up the video stream and starts displaying it, without further interaction from the user. If the user is done using the camera unit and it gets shut-off, see subsection 4.2, the stream presentation shall gracefully turn off.

The implementation of the script does the following things, as can be seen in Figure 4: First, the script gets only executed if the DRS user is logged in. If that is the case, it enters an infinite while-loop, which gets executed every three seconds. The display/control unit starts pinging the camera unit with the *ping* command. If the command is successful, the script enters the next stage: It tries to connect to the RTSP/RTP socket via *netcat*, to verify that the stream is actually available. If this operation is also successful, the scripts checks if the stream application is already running and displaying the video stream. If that is the case, it goes back to the beginning of the while-loop. If the stream isn't already running, it will start the stream application and the stream gets displayed. If the connection attempt with netcat or the ping command are not successful, the script assumes that the camera unit is not operating or isn't available. The script will then check if the stream application is currently running. If that is the case, the stream application will be gracefully shut-down and the script will return to the start of the while-loop. If not, then there is nothing to do and the cycle also begins from the start.

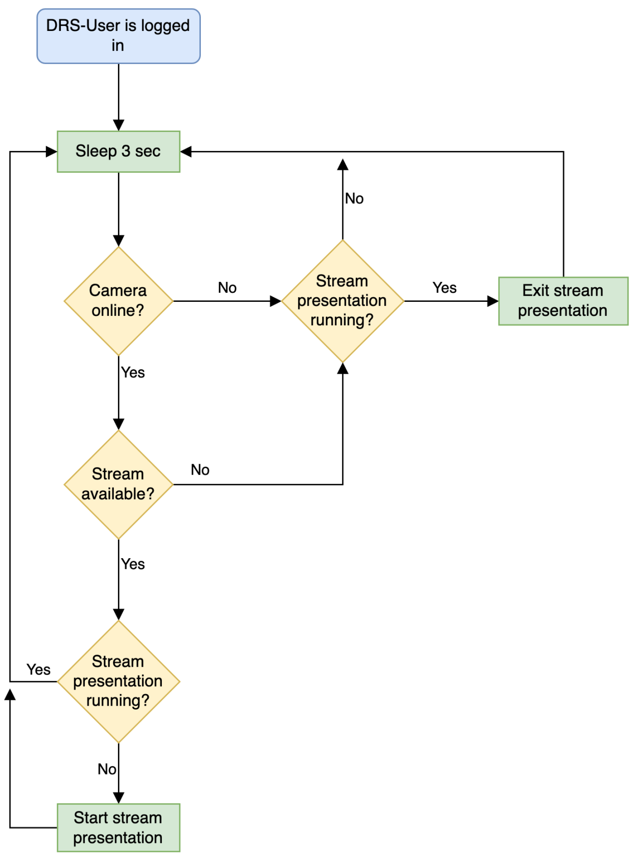



**Fig. 4.** Activity diagram of the automated start/stop control of the video stream

### 4.2 Automated shut-off of camera unit

If the user is finished using the DRS, the camera unit just needs to be placed somewhere stationary and it shall turn off automatically after a certain amount of time to preserve battery life.

To achieve this behaviour, the stream application has a built-in functionality, which tries to recognize movement in the video stream, which also can be seen in Figure 5. It implements a naive approach by calculating the mean value of the current frame and comparing it to the last frame. If the value is less than a certain threshold, the script assumes that there was no movement between the two frames and thus the camera is not moving or displaying new content. It then increases a counter, which constantly gets checked against another threshold: If the counter is greater than this threshold, the camera is probably not getting used anymore and the shut-down procedure will start. The shut-down procedure will start a timer and will embed a warning in the video stream, that the stream will soon stop. After a certain amount of time the display/control unit will connect to the camera unit via SSH and shut-down the unit. If, however, the difference in the mean value of the two frames is greater than the threshold, the counter will reset to zero. That results in a very sensitive movement detection, because only one major changed frame will reset the counter to zero. The current configuration of the thresholds results in a ten minute window before the shut-off procedure will shut-off of the camera unit.

This mean calculating and difference comparison method is very light-weight on the CPU, but does not offer the best recognition. In a later iteration of the DRS, a more advanced method to recognize movement in the stream should be used.

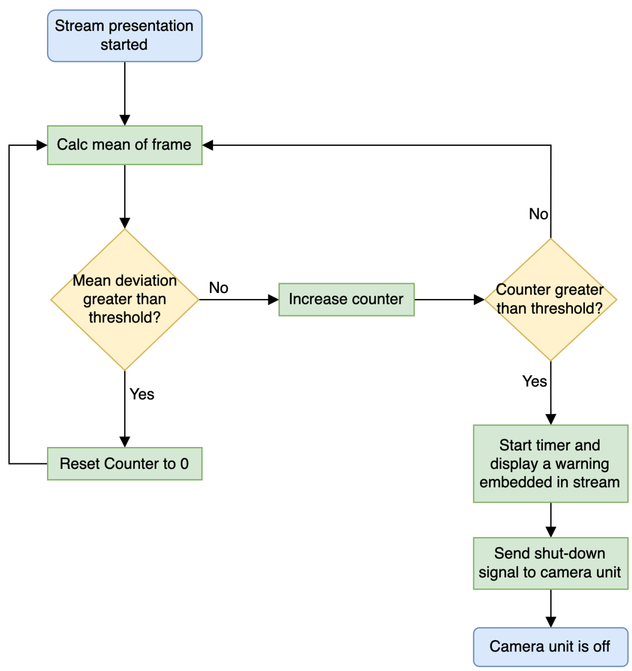



**Fig. 5.** Activity diagram of the automated shut-off functionality of the camera unit

## 5 Automated set-up of components

Another aspect of automating as many things as possible is to also automate the set-up process of the camera and display/control unit, to be able to quickly change the used hardware. To realize that goal, the automation software *Ansible* is used, where you can specify certain actions on a target device, which then get automatically executed.

The following actions are performed automatically:

- Install all necessary software and libraries on the devices, which are needed for operation
- Copy the scripts, mentioned in section 4, to the devices, respectively
- Configures the auto-start of these scripts at boot time

# 6 Outlook

The whole project is still a work in progress and misses some essential parts. Following bullet-points need to be solved in the future:

- A fully operational stand for the camera unit has to be designed, where a text-source e.g. a book can be placed beneath.
- A fully closed housing for the camera unit and the powerbank has to be designed and built. It shall protect the camera unit and also keep the components in place.
- A better camera has to be found, to further improve the OCR. Fully operational mechanical auto-focus would be ideal, as well. Also the use of a light source should be analyzed. A good candidate is the Raspberry Pi Camera Module 3 with included Auto-Focus and HDR.
- Implementing the TTS pipeline and researching a convenient activation method for it
- Real-World test for fine-tuning parameters

# IDOVIR – Infrastructure for Documentation of Virtual Reconstructions

Jonas Bruschke[1], Markus Wacker[1], Marc Grellert[2], Daniel Beck[3], and Wolfgang Stille[3]

[1] Department of Computer Graphics, University of Applied Sciences Dresden, Germany
[2] Department of Architecture, TU Darmstadt, Germany
[3] hessian.AI, Darmstadt, Germany

**Abstract.** Virtual reconstructions have become a common tool for research and communication in the field of architectural and urban history studies. They often visualize architecture that no longer exists, but also never realized construction plans. In general, such reconstructions are created based on various sources that are often ambiguous or incomplete. Though, the intellectual, decision-making processes that underpin these reconstructions ususally remain intransparent and the knowledge and findings are in danger of being lost. For the scientific community however, it is important that the processes and knowledge behind a reconstruction is publicly accessible and documented in a transparent, comprehensible, and sustainable manner. IDOVIR aims to overcome these shortcomings by establishing a free, accessible, web-based online platform, with which the documentation of virtual reconstructions and the decision-making processes behind can be accomplished in a supportive way and will be publicly available.

## 1 Introduction

Virtual reconstructions have become a common tool for research and communication in the field of architectural and urban history studies [5]. They often visualize architecture that no longer exists, but also never realized construction plans as well as former states of an existing building or even whole cities. In general, such reconstructions are created by humans by interpreting various sources. The source material, however, is usually not coherent or complete. Decisions have to been made on parts where information is ambiguous or missing. New knowledge about the reconstructed object can be gained from this process [3]. The audience, however, usually only sees the final results. The intellectual, decision-making processes that underpin these digital reconstructions remain intransparent and the knowledge and findings are in danger of being lost. For the scientific community however, it is important that the processes and knowledge behind a reconstruction is publicly accessible and documented in a transparent, comprehensible, and sustainable manner. Only if this is given, a proper scientific evaluation and discussion can be possible on the results as well as findings can be continued to use in further research.

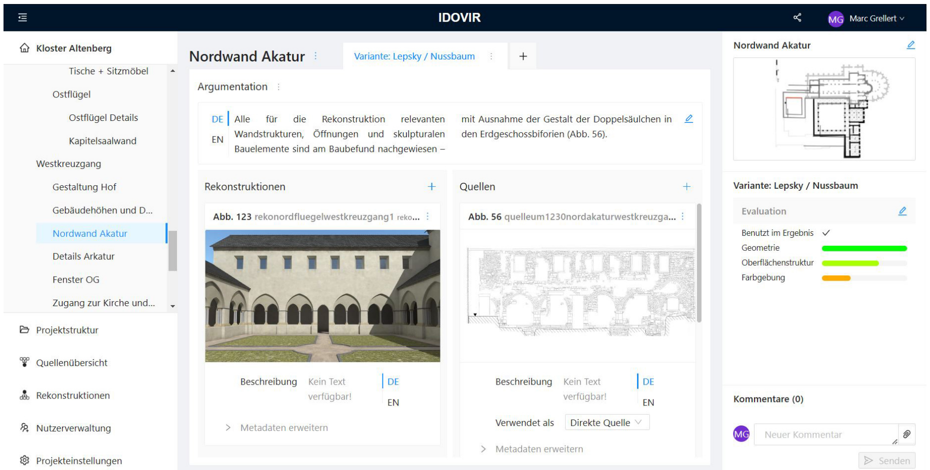



**Fig. 1.** Graphical user interface of the IDOVIR web application outlining the argumentation connecting a reconstruction with sources.

However, the current situation is that a proper, scientifically valid documentation is usually not done at all or only insufficiently [6]. The community has long been aware of this problem and also published theoretical policy papers such as the London Charter [2] that formulate principles regarding sustainability, verifiability, and knowledge preservation. There are many approaches concerning the documentation of individual projects or single aspects of documentation such as the visualization of reliability [7]. Nevertheless, documentation is still an exception. On the one side, a proper documentation is expensive and there are usually not enough human and financial resources [5]. Additionally, it is generally not explicitly demanded by the funding bodies. On the other side, there is no common agreement on the standards, structure, and content of such documentations as well as best practices.

## 2 Towards a common documentation tool

IDOVIR aims to establish a free, accessible, web-based online platform[1], with which the documentation of virtual reconstructions and the decision-making processes behind can be accomplished and will be publicly available. For long-term accessibility and maintenance, the application is embedded in the system landscape fo the University and State Library Darmstadt. The project is based on preliminary work that aimed to ease and support the documentation process. This includes the online tool ScieDoc[2] by TU Darmstadt and the DokuVis[3]

[1] `https://idovir.com/`
[2] Scientific Documentation for Decision, `http://sciedoc.org/`
[3] Documentation Tool for Digital 3D Reconstructions, `http://dokuvis.org/`

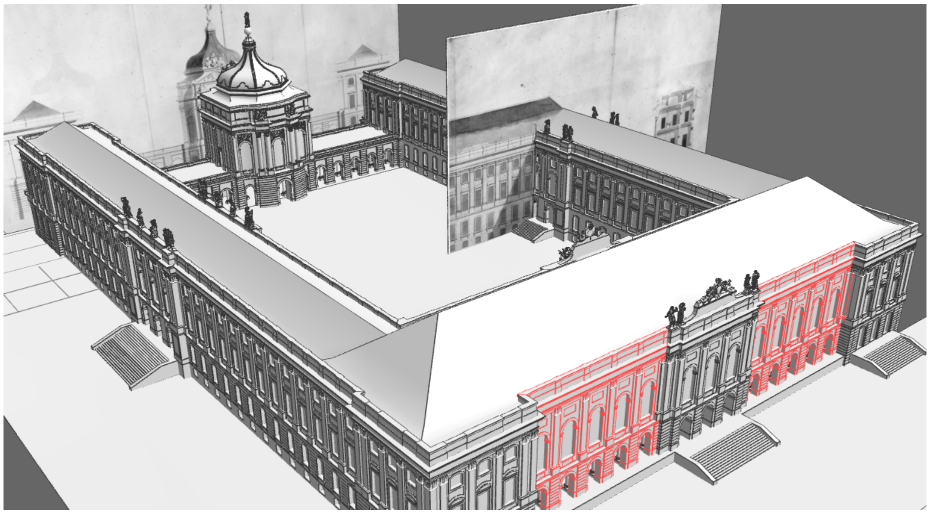


**Fig. 2.** For evaluation, source material like construction plans can be visualized together with the 3D model of the reconstructed object and interactively compared against each other.

prototype by the University of Applied Sciences Dresden. The advantages and synergies of both systems are combined in a joint platform.

There are already many projects that have been documented with ScieDoc. The focus lies on a minimum documentation that covers only the most basic information to comprehend the reconstruction results [4]. In this regard, multiple areas can be defined to subdivide the object of reconstruction. For each area, one or more variants can be created (multiple variants may be necessary if the source material is ambiguous or contradictory). A variant consists of the core documentation: the triple of images of the reconstruction, the source material, and a textual argumentation, i.e., a qualitative analysis that links the reconstruction with the sources on which it is based (cf. Fig. 1). The system was designed with only little features that are, though, quick to master and lowers the acceptance threshold. These basic features have already been implemented in IDOVIR with some improvements and additions such as hierarchical ordering of the subdivisions (i.e., project structure), semantic differentiation between areas and timesteps, evaluation of variants, and proper management of uploaded media. The objective to provide an intuitive workflow and a low-threshold entry point remains at high priority. This becomes more challenging as more features are getting integrated.

DokuVis took a different approach. Here, the 3D model, which is usually the main outcome of the reconstruction process, plays a central role in visualizing information and being an object of interaction [1]. Sources can be directly compared against the model within the 3D viewport supported by various rendering styles (Fig. 2). Additional information can be attached to the place of interest within the model. The scope of DokuVis comprises not only the documentation, but also the collaboration aspect envisaging a usage of the tool right from the start accompanying the whole reconstruction process. The idea is that if the tool helps in managing and supporting this process, the documentation is done almost incidentally and not much extra work is required. In this regard, features like project management and a simple version control of 3D models have been prototypically implemented next to communication features like comments.

IDOVIR wants to combine both approaches: On the one side, there is this minimalistic approach of ScieDoc with its low-threshold entry point. On the other side, the feature-rich approach of DokuVis offers advanced documentation and collaboration, but also faces challenges regarding usability. Thus, within the tool, users should be able to choose different views depending on their expertise and data situation: (a) input option with minimal effort, (b) guided data input with user-defined fields, (c) two-dimensional representation of virtual reconstructions in the form of renderings, (d) three-dimensional representation of reconstructions in the form of an interactive 3D model, and (e) various evaluation tools, e.g., measuring distances in the model. Additional feature requests by the target group go beyond the preliminary work and have already been partially implemented. This includes export as consolidated well-laid-out PDF documents, georeferencing, and annotations in different kinds of documents.

## 3 Outlook

A first version of IDOVIR is already online comprising the basic documentation features with 2D images. Many project partners, who were using one of the predecessor tools, are now testing the application. The integration of 3D models is currently under development. There is still potential to further support the data input to keep the work load to a minimum. Hence, some workflows can be assisted by (partially) automated processes. For example, metadata about the sources and the reconstructed object may be queried automatically from other instances such as Wikidata, the German Digital Library, or other repositories.

In general, projects can be set to be public, so they are open for discussion for other researchers. Many of the sources were retrieved from archives and libraries, so they actually hold the rights for publication. An open issue, however, is how to deal with uploaded material with unknown copyright information.

## Acknowledgment

This work has been funded by the Deutsche Forschungsgemeinschaft (DFG), grant identifier 460771453.

## References

1. Bruschke, J., Wacker, M.: Simplifying documentation of digital reconstruction processes. In: Münster, S., Pfarr-Harfst, M., Kuroczyński, P., Ioannides, M. (eds.) 3D

Research Challenges in Cultural Heritage II: How to Manage Data and Knowledge Related to Interpretative Digital 3D Reconstructions of Cultural Heritage. pp. 256–271. Springer International Publishing, Cham (2016)

2. Denard, H.: The london charter for the computer-based visualisation of cultural heritage (2009), http://www.londoncharter.org
3. Echtenacher, G.: Wissenschaftliche Erkenntnisse durch manuelles Konstruieren von 3D-Modellen. In: Heine, K., Rheidt, K., Henze, F., Riedel, A. (eds.) Von Handaufmaß bis High Tech III. pp. 49–57. Verlag Philipp von Zabern, Darmstadt/Mainz (2011)
4. Grellert, M., Pfarr-Harfst, M.: Die rekonstruktion-argument-methode – minimaler dokumentationsstandard im kontext digitaler rekonstruktionen. In: Kuroczyński, P., Pfarr-Harfst, M., Münster, S. (eds.) Der Modelle Tugend 2.0: Digitale 3D-Rekonstruktion als virtueller Raum der architekturhistorischen Forschung. p. 264–280. arthistoricum.net, Heidelberg (2019)
5. Münster, S.: Interdisziplinäre Kooperation bei der Erstellung geschichtswissenschaftlicher 3D-Modelle. Springer VS, Wiesbaden (2016)
6. Pfarr, M.: Dokumentationssystem für Digitale Rekonstruktionen am Beispiel der Grabanlage Zhaoling, Provinz Shaanxi, China. Dissertation, TU Darmstadt (2010)
7. Wacker, M., Bruschke, J.: Dokumentation von digitalen rekonstruktionsprojekten. In: Kuroczyński, P., Pfarr-Harfst, M., Münster, S. (eds.) Der Modelle Tugend 2.0: Digitale 3D-Rekonstruktion als virtueller Raum der architekturhistorischen Forschung. p. 282–294. Computing in Art and Architecture, arthistoricum.net, Heidelberg (2019)

# Tracking multiple VR users in a shared physical space

Isabella Skořepová

CTU Prague

**Abstract.** Multi-user experiences in shared physical spaces are a staple of VR arcades, but creating them using consumer-grade hardware is still complicated. This text will discuss options on how to implement them and describe the author's work on one such implementation, which attempts to simplify the creation of such experiences. In addition, various technical aspects of such implementation will be discussed, along with some required non-technical considerations. Implementation is done for the Unity engine using the OpenXR plugin in Rust programming language and should work across standalone and PC-based headsets. It should be possible to port the work to other game engines, but that is out of the scope of this work

## 1 Introduction

Imagine you have an idea of a perfect virtual reality (VR) experience. I could be a multiplayer VR arena shooter à la laser tag, an exposition in an art gallery, or an escape room. And you want it be for multiple people in the same room - in a shared physical space.

How would you create such an experience? What problems do you need to tackle before making your idea a reality?

Let's break it down into different parts: You have multiple users, in the same space, each wearing their own headset. Each headset has its own computer[1] and depending on headset possibly has its own tracking system with separate origins. Author of such VR experience has to make all of that talk to each other and align, on top of creating the VR experience itself.

Local-network real-time multiplayer is well-researched and well-understood problem, which this article will not delve into. What's of interest to us is how to make the coordinate systems align and the specific implementation for recent VR headsets.

## 2 Alternatives

But first, before we delve into the specific solution chosen for my work, I'd like to talk about other possible solutions and their problems.

[1] It is either attached to in case of PCVR, or is it's own computer in the case of standalone headsets

One such solution is to use a system which has shared global coordinates. One such system is OptiTrack. This is a well-known altenative, which unfortunately is quite expensive. Another option is to only use headsets which use lighthouse-based tracking as those can also be shared between multiple headsets (as opposed to oculus now discontinued constellation tracking) and has basestations whose position it can reference.

I decided to use neither and to come up with a system, which would be possible to use across various devices, including devices using different tracking system at the same time. And which would work with just the bare devices themselves.

## 3 Solution

The solution is to use a software-based calibration methods. It is implemented as OpenXR layer, so it stands between game engine and OpenXR[2] runtime. It makes the engine think that controllers from other headsets are also connected to local headset. This is already enough to implement some most basic experiences. And to make it possible to make more complicated experiences, I add a simple extension which allows the application to add specific support for remote devices.

## 4 Calibration

The most interesting part, as alluded to previously, is the calibration. I came up with two methods, which I already implemented in prototypes and both of them show great promise in simple testing.

First calibration method is what I currently call "tower of doom". For this we take advantage of the fact, that the VR headsets already calibrate down vector and position of floor. The calibration therefore becomes a 2D problem. All headsets I have access to for testing have a property that they can be stacked on top of each other. Therefore we can simultaneous calibrate multiple headsets by putting them on top of each other and making the direction they are looking at the same. This in preliminary testing works well for multiple headsets of the same model, but also works surprisingly well for different models of headsets.

Second calibration method is more elaborate, but also more precise. It was originally used in OpenVR Space Calibrator, but I made a few improvements to the original method.[3] It works by holding controllers from multiple different headset systems in tight grip (either using your hand, or a jig) and making infinity symbol with them. During this motion samples are collected, which are then fed to an algorithm and this results in position and rotation offsets. To calibrate more then two headsets you either hold all of them at once, or calibrate

[2] OpenXR is a standard supported by all current VR headsets. Not all engines currently use OpenXR on all platforms, but at the time of writing, it seems to be where the industry is going in the near to medium term future.

[3] If those improvements improve any key metric is yet to be determined.

each headset with once chosen central headset. Alternatively, you could devise more complicated schemes, but the error compounds, so it is best to keep the calibration path between any two headsets as short as possible.

## 5 Future work

There are many aspects which affect such systems in different ways. Most of them I couldn't get into in this short text. My goal is to finish a working system and use it explore those aspects and document them.

## Acknowledgement

This work has been supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS22/173/OHK3/3T/13 (Research of Modern Computer Graphics Methods 2022-2024).

# Towards Aesthetics of Subjectivity in InfoVis

Elena Stoll

Department of Visual Engineering, University of Applied Sciences Dresden,
Friedrich-List-Platz 1, 01069 Dresden, Germany

**Abstract.** Aesthetics is a topic of rising interest in information visualization. User studies have demonstrated positive effects on interest, memorability, performance, or enjoyment. In contrast to research on aesthetics in other disciplines, the notion of aesthetics in information visualization is not only blurry but somewhat objectified, i.e. tied to properties of the visualization itself. Thus, subjective aspects which are known to be a driving factor of aesthetic experience are often neglected. In this paper, research on aesthetic experience and measurement tools is reviewed regarding the applicability to information visualization. Directions for further research are outlined, particularly arguing for the development of an aesthetic theory and measurement methods for information visualization that embrace subjective experiences.

## 1 Introduction

Information visualization (InfoVis) has been described as the interactive communication of abstract data, concepts, or processes to amplify human cognition [23, 9]. Most recently, the discussion of beauty and aesthetics has gained attention in the visualization community [41, 7], marking a shift away from functional and data-centric approaches, by which visualizations are created traditionally [11, 1]. User studies have shown that aesthetics can positively affect the viewers' interest, memorability, performance, or joy of use [8]. Research predominantly focuses on the aesthetics of single aspects of visualizations, such as color, layout, or diagram type [44, 40]. In general, aesthetics in InfoVis are mainly seen as something that stems from objective properties of a visualization, for example symmetry, harmony, or coloration. This partly contradicts findings from other disciplines such as experimental aesthetics, cognitive psychology, or neuroaesthetics. Research in these areas suggests that subject-related and contextual factors play a central role in how humans experience aesthetics [42]. Consequently, there seems to be a need to broaden the view of aesthetics in InfoVis and include subject-related factors through modeling visualization processing from the viewer's perspective and developing methods to capture individual experiences of aesthetics.

The remainder of the paper offers a first approach to the described problem. In a cross-domain overview of related research, the concept of aesthetics per se as well as in relation to InfoVis are illuminated, and an introduction to perception and measurement of aesthetic experience is given. Results are then synthetized and questions for further research are outlined. Finally, a summary and an outlook are given.

## 2 Related Work

*The Concept of Aesthetics.* The origins of aesthetics (after ancient Greek aísthēsis = sensual perception) date back to ancient philosophy of Plato and Aristotle and Alexander Gottlieb Baumgarten, who coined the term for a "science of sensual cognition" in his publication Aesthetica in 1750 [3]. Three well-known, traditional philosophical definitions of aesthetics are "theory of art," "theory of the beautiful," and "theory of sensuous cognition." These definitions have been found insufficient by Reicher, who instead proposes a philosophical explication stating that aesthetics are an interplay of the following elements: (1) aesthetic properties (e.g., beauty, ugliness, harmony, sublimity, and many others), (2) aesthetic objects (including immaterial objects such as ideas or mental images, which have, also, but not exclusively, aesthetic properties), and (3) aesthetic experiences (in which aesthetic properties are captured) [36]. Hence, aesthetics is not a property of objects, but an individual experience triggered by the perception and mental processing of aesthetic properties of objects, situations, etc. This distinction might help to frame publications in which the terms beauty and aesthetics are used interchangeably.

*Aesthetics in InfoVis.* Research is mainly concerned with the importance of aesthetics for the visualization domain [7, 41] or the aesthetics of specific aspects of visualizations such as color [44], shape [8], layout [45], or diagram type [40, 4]. Relatively few authors focus on subjective factors of aesthetics, such as inter-individual differences in aesthetic evaluation of graphical patterns [21], the relevance of subjective perception in dashboard design [20], individual preferences in graph design [26], or the relationship between aesthetics and usability of visualizations [10]. While these approaches take into account the variety of inter-individual aesthetic perception, they remain entrenched in single aspects. An overarching theory visualization aesthetics, which, for example, picks up the idea of aesthetics as a holistic experience, is missing so far. Another open issue is that there seems to be some ambiguity in the use of the terms aesthetics and beauty, which are used interchangeably, inconsistently, and vaguely.

*Aesthetic Experience and Emotions.* In experimental aesthetics, the processing of stimuli and aesthetic experiences are investigated empirically on the physiological, mental, and behavioral levels. A central, yet unresolved question is to what extent aesthetic perception is objective-static or subjective-flexible. Many attempts have been made to establish universally applicable rules of aesthetics such as "round is beautiful" [2]. According to the aesthetic principle of association of Fechner, aesthetic perception is primarily influenced by memories and associations, i.e. the individual learning history: "We believe that we are able to read the entire beauty off a beautiful arm, without suspecting that we read the most of it into the same." [31, p. 6]. Aesthetic evaluations are also subject to dynamic changes, as demonstrated by studies using semantic or contextual priming [13, 22].

Neuroaesthetics uses modern imaging techniques to study brain activities linked to aesthetic experiences. For example, the activation of brain regions for motor function, imagination, memory, and emotion has been visualized [12]. Currently, aesthetic emotions are also being explored and how they differ from other, basic emotions. Aesthetic emotions seem to cover a broad spectrum, including negative emotions [39]. A key difference that has emerged is that aesthetic emotions often lack a pragmatic motivation for action or goal-directedness, as they seem to be strongly related to the reward system for liking and enjoyment (liking system) rather than triggering a desire for satisfaction of needs (wanting system) [27].

*Modeling Aesthetics.* In perceptual psychology or (affective) neuroscience, models of aesthetic experience have been proposed [15, 25, 32, 35, 43]. Usually, the models combine bottom-up sensory processing and top-town mental processing, which includes, for example, domain knowledge, or individual experiences. Starting with a global, non-specific exploration, a deep cognitive and emotional aesthetic evaluation gradually emerges throughout the process [25]. The aesthetic experience is shaped by affective and cognitive processing with feedback and feedforward mechanisms, influencing each other mutually. However, it remains unclear to what extent consciousness plays a role in the context of aesthetic experience. Also central to aesthetic models is the Fluency Theory [15, 34]. Broadly speaking, objects which the brain can process more easily (”fluently”) due to certain properties receive higher or more extreme aesthetic ratings. A sudden increase in fluency, combined with the perceived certainty of having gained new insights or knowledge, has been described as the Aesthetic Aha Effect [29].

*Measuring Aesthetics.* In ”Aesthetic Measure” (1933), mathematician George D. Birkhoff suggests an aesthetic measure as relation of order to complexity (M=O/C) [5]. Birkhoff’s idea was often taken up, for example in the ”Pattern Measure” for simple patterns or in information theory [14, 24]. In the field of computational aesthetics, rule-based approaches or methods of machine learning, in particular neural networks, are used to predict an aesthetic image value based on objective image features or to generate aesthetic images [17, 18]. A computer model of visual human stimulus processing are trained predict subjective value judgments [6]. To quantify of visual aesthetics, a number of measuring tools has been developed. The Aesthetic Measurement Application (AMA) [28] captures the aesthetics of web interfaces. The Visual Aesthetic Sensitivity Test-Revised (VAST-R) [30] is based on Eysenck’s idea of a general factor for visual aesthetic evaluation (so-called ”T” for good taste). With the help of the Aesthetic Emotions Scale (AESTHEMOS) [38], a range of aesthetic emotions and their activating or sedative effect is detected. With Athec [33], a ready-to-use Python library is available to measure the aesthetics of visual media. BeauVis [16] is a scale for measuring aesthetic pleasure of visual representations. A dedicated instrument to capture also the subjectively experienced aesthetics of InfoVis not yet exists.

## 3 Research Roadmap

Based on the interdisciplinary foray into aesthetics, it can be stated that aesthetics is a multi-level experience with both objective and subjective components. The following diagram summarizes various models of aesthetic experience, depicting how aesthetic judgments and emotions can arise from visual stimuli.

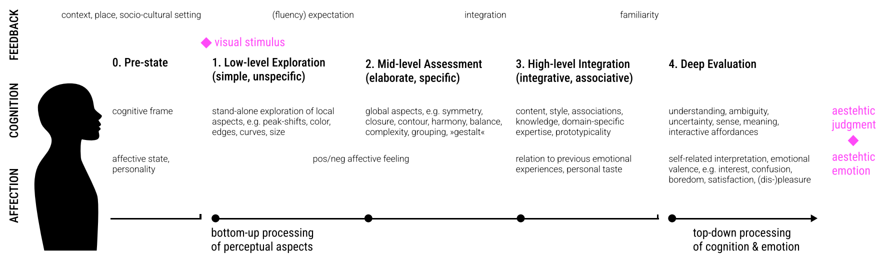



**Fig. 1.** Schematic representation of aesthetic experience, based on [15, 25, 32, 35, 43].

Two main questions evolved that could guide further research:

1. Which aesthetic models or frameworks exist that support the development of an aesthetic theory of InfoVis?
2. What could an aesthetic metric look like that captures individual aesthetic experiences related to information visualizations?

A more comprehensive, interdisciplinary literature review is necessary to answer the questions. The tentative model can then be improved and transferred to the field of InfoVis. In order to augment the metrics by subjective aspects, experimental studies can be carried out on the model of psychology or neuroscience. For both questions, a major challenge is to balance objective and subjective aspects of aesthetics, i.e. to find the middle ground.

## 4 Conclusion and Outlook

In this article, the tension between InfoVis and aesthetics, including subjective aspects, was examined. Based on an overview of selected literature on aesthetics, directions for further research were indicated. The topic can also be considered on a larger scale. Technological developments such as automation, massive computing capacities and data availability, and machine learning can open up new ways to address complex challenges – challenges to which aesthetic issues most likely belong. Existing attempts to automate the generation of visualizations [37, 19] could serve as a technical prerequisite towards the creation of context-dependent visualizations that reflect individual needs and preferences.

**Acknowledgements** This work has been supported by the European Regional Development Fund and the Free State of Saxony (project no. 100602780). Thanks are due to WaPro project supervisor Prof. Dr.-Ing. Dietrich Kammer for his comments on the paper draft.

## References

1. Aigner, W.: InfoVis:Wiki: Visualization Pipeline, DOI: `https://infovis-wiki.net/wiki/Visualization_Pipeline`, visited 2023-02-17
2. Bar, M., Neta, M.: Humans prefer curved visual objects. Psychological Science 17(8), 645–648 (2006)
3. Baumgarten, A.G.: Aesthetica. J.C. Kleyb, Frankfurt a.d. Oder (1750)
4. Bennett, C., Ryall, J., Spalteholz, L., Gooch, A.: The Aesthetics of Graph Visualization. Computational Aesthetics in Graphics Visualization (2007), DOI: 10.2312/COMPAESTH/COMPAESTH07/057-064
5. Birkhoff, G.D.: Aesthetic Measure. Harvard University Press, Cambridge, MA and London, England (1933), DOI: doi:10.4159/harvard.9780674734470
6. Brielmann, A.A., Dayan, P.: A computational model of aesthetic value. Psychological Review 129(6), 1319–1337 (2022), DOI: 10.1037/rev0000337
7. Brinch, S.: What we talk about when we talk about beautiful data visualizations. In: Engebretsen, M., Kennedy, H. (eds.) Data Visualization in Society, pp. 259–276. Amsterdam University Press, Amsterdam (2020), DOI: doi:10.1515/9789048543137-020
8. Carbon, C.C., Mchedlidze, T., Raab, M.H., Wächter, H.: The power of shape: How shape of node-link diagrams impacts aesthetic appreciation and triggers interest. i-Perception 9(5), 2041669518796851 (2018), DOI: 10.1177/2041669518796851
9. Card, S.K., Mackinlay, J.D., Shneiderman, B.: Readings in information visualization: using vision to think. The Morgan Kaufmann series in interactive technologies, Morgan Kaufmann Publishers, San Francisco (1999)
10. Cawthon, N., Moere, A.V.: The effect of aesthetic on the usability of data visualization. In: 2007 11th International Conference Information Visualization (IV '07). pp. 637–648 (2007), DOI: 10.1109/IV.2007.147
11. Chi, E.: A taxonomy of visualization techniques using the data state reference model. In: IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings. pp. 69–75 (2000), 10.1109/INFVIS.2000.885092
12. Di Dio, C., Gallese, V.: Moving Toward Emotions in the Aesthetic Experience. In: Brain, Beauty, and Art, pp. 22–26. Oxford University Press (Jan 2022), DOI: 10.1093/oso/9780197513620.003.0005
13. Faerber, S.J., Leder, H., Gerger, G., Carbon, C.C.: Priming semantic concepts affects the dynamics of aesthetic appreciation. Acta Psychologica 135(2), 191–200 (2010), DOI: 10.1016/j.actpsy.2010.06.006
14. Filonik, D., Baur, D.: Measuring aesthetics for information visualization. In: 2009 13th International Conference Information Visualisation. pp. 579–584 (2009), 10.1109/IV.2009.94
15. Graf, L.K.M., Landwehr, J.R.: A dual-process perspective on fluency-based aesthetics: The pleasure-interest model of aesthetic liking. Personality and Social Psychology Review 19(4), 395–410 (2015), DOI: 10.1177/1088868315574978

16. He, T., Isenberg, P., Dachselt, R., Isenberg, T.: Beauvis: A validated scale for measuring the aesthetic pleasure of visual representations. IEEE Transactions on Visualization and Computer Graphics PP, 1–11 (01 2022), DOI: 10.1109/TVCG.2022.3209390
17. den Heijer, E., Eiben, A.E.: Investigating aesthetic measures for unsupervised evolutionary art. Swarm Evol. Comput. 16, 52–68 (2014)
18. Hoenig, F.: Defining computational aesthetics. In: Proceedings of the First Eurographics Conference on Computational Aesthetics in Graphics, Visualization, and Imaging. pp. 13–18 (01 2005), DOI: 10.2312/COMPAESTH/COMPAESTH05/013-018
19. Hu, K.Z., Bakker, M.A., Li, S., Kraska, T., Hidalgo, C.A.: Vizml: A machine learning approach to visualization recommendation. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (2018), DOI: 10.1145/3290605.3300358
20. Hynek, J.: Impact of Subjective Visual Perception on Automatic Evaluation of Dashboard Design. Ph.D. thesis, Brno University of Technology (2019), DOI: 10.13140/RG.2.2.28372.96649
21. Jacobsen, T., Höfel, L.: Aesthetic judgments of novel graphic patterns: Analyses of individual judgments. Perceptual and Motor Skills 95(3), 755–766 (2002), DOI: 10.2466/pms.2002.95.3.755
22. Jakesch, M., Zachhuber, M., Leder, H., Spingler, M., Carbon, C.C.: Scenario-based touching: on the influence of top-down processes on tactile and visual appreciation. Research in Engineering Design 22(3), 143–152 (2011), DOI: 10.1007/s00163-010-0102-5
23. Keim, D., Mansmann, F., Schneidewind, J., Ziegler, H.: Challenges in Visual Data Analysis. In: Tenth International Conference on Information Visualisation (IV'06). pp. 9–16. IEEE, London, England (2006), DOI: 10.1109/IV.2006.31
24. Khalili, A.M., Bouchachia, A.: An information theory approach to aesthetic assessment of visual patterns. Entropy 23 (2021)
25. Leder, H., Nadal, M.: Ten years of a model of aesthetic appreciation and aesthetic judgments : The aesthetic episode - developments and challenges in empirical aesthetics. British Journal of Psychology 105(4), 443–464 (2014), DOI: 10.1111/bjop.12084
26. McGurgan, K., Fedoroksaya, E., Sutton, T., Herbert, A.: Graph design: The data-ink ratio and expert users:. In: Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. pp. 188–194. SCITEPRESS - Science and Technology Publications (2021), DOI: 10.5220/0010263801880194
27. Menninghaus, W., Wagner, V., Wassiliwizky, E., Schindler, I., Hanich, J., Jacobsen, T., Koelsch, S.: What are aesthetic emotions? Psychological Review 126(2), 171–195 (Mar 2019), DOI: 10.1037/rev0000135
28. Mohamad Zain, J., Tey, M., Soon, G.: Using aesthetic measurement application (ama) to measure aesthetics of web page interfaces. In: 2008 Fourth International Conference on Natural Computation. pp. 96 – 100 (11 2008), DOI: 10.1109/ICNC.2008.764
29. Muth, C., Carbon, C.C.: The aesthetic aha: On the pleasure of having insights into gestalt. Acta Psychologica 144(1), 25–30 (2013), DOI: 10.1016/j.actpsy.2013.05.001
30. Myszkowski, N., Storme, M.: Measuring "good taste" with the visual aesthetic sensitivity test-revised (vast-r). Personality and Individual Differences 117, 91–100 (05 2017), DOI: 10.1016/j.paid.2017.05.041

31. Ortlieb, S.A., Kügel, W.A., Carbon, C.C.: Fechner (1866): The aesthetic association principle—a commented translation. i-Perception 11(3) (2020), DOI: 10.1177/2041669520920309
32. Pelowski, M., Markey, P.S., Forster, M., Gerger, G., Leder, H.: Move me, astonish me… delight my eyes and brain: The vienna integrated model of top-down and bottom-up processes in art perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates. Physics of Life Reviews 21, 80–125 (2017), DOI: 10.1016/j.plrev.2017.02.003
33. Peng, Y.: Athec: A python library for computational aesthetic analysis of visual media in social science research. Computational Communication Research 4, 323–349 (2022), DOI: 10.5117/CCR2022.1.009.PENG
34. Reber, R., Schwarz, N., Winkielman, P.: Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? Personality and social psychology review : an official journal of the Society for Personality and Social Psychology, Inc 8, 364–82 (02 2004), DOI: 10.1207/s15327957pspr0804_3
35. Redies, C.: Combining universal beauty and cultural context in a unifying model of visual aesthetic experience. Frontiers in Human Neuroscience 09 (2015), DOI: 10.3389/fnhum.2015.00218
36. Reicher, M.E.: Einführung in die philosophische Ästhetik. Einführung Philosophie, WBG (Wissenschaftliche Buchgesellschaft), Darmstadt, 3rd revised edition edn. (2015)
37. Satyanarayan, A., Moritz, D., Wongsuphasawat, K., Heer, J.: Vega-lite: A grammar of interactive graphics. IEEE Transactions on Visualization and Computer Graphics 23(1), 341–350 (2017), 10.1109/TVCG.2016.2599030
38. Schindler, I., Hosoya, G., Menninghaus, W., Beermann, U., Wagner, V., Eid, M., Scherer, K.R.: Measuring aesthetic emotions: A review of the literature and a new assessment tool. PLoS ONE 12 (2017), DOI: 10.1371/journal.pone.0178899
39. Silvia, P.J.: Looking past pleasure: Anger, confusion, disgust, pride, surprise, and other unusual aesthetic emotions. Psychology of Aesthetics, Creativity, and the Arts 3(1), 48–51 (Feb 2009), DOI: 10.1037/a0014632
40. Strunge Mathiesen, S., Schulz, H.J.: Aesthetics and ordering in stacked area charts. In: Diagrammatic Representation and Inference: 12th International Conference, Diagrams 2021, Virtual, September 28–30, 2021, Proceedings. p. 3–19. Springer-Verlag, Berlin, Heidelberg (2021), DOI: 10.1007/978-3-030-86062-2_1
41. Vande Moere, A., Schmidt, C., Lima, M.S.: Aesthetics in information visualization (2009)
42. Vessel, E.A., Starr, G.G.: Imaging the Subjective. In: Brain, Beauty, and Art: Essays Bringing Neuroaesthetics into Focus. Oxford University Press (01 2022), DOI: 10.1093/oso/9780197513620.003.0024
43. Xenakis, I., Arnellos, A.: Aesthetic perception and its minimal content: a naturalistic perspective. Frontiers in Psychology 5 (2014), DOI: 10.3389/fpsyg.2014.01038
44. Xie, Y., Hua, L., Chen, R., Lin, T., Tang, N.: Using color to improve the discrimination and aesthetics of treemaps. In: 2016 Nicograph International (NicoInt). pp. 143–143 (2016), 10.1109/NicoInt.2016.36
45. Zheng, Q., Li, Z., Bargteil, A.: Learning aesthetic layouts via visual guidance (2021), DOI: 10.48550/ARXIV.2107.06262

# VentConnect: live to life and the octopus in the hospital server room

Jan Jirman, Miroslav Macík

Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic

**Abstract.** In the ICU environment, the parameters of various instruments connected to the patients must be set correctly, especially in the medium to long-term processes. The patient's state develops over time, and these parameters must be re-adjusted. Our project provides a streamlined process for checking the parameters and the state of the patient by presenting all the necessary information to the physicians remotely – without needing to visit each bed every single time. This has proved to be particularly important during the COVID-19 pandemic. During the pandemic, we have focused mainly on mechanical lung ventilation; however, the system can also be used for other purposes or instruments.

## 1 Introduction

VentConnect is a project that started as a response to the COVID-19 pandemic. Specifically, it addresses certain challenges that the hospital staff had with mechanical lung ventilation.

Mechanical lung ventilation is a medium-term process, meaning the physician sets up the ventilator once and then leaves it. Throughout the day, as the patient's state changes, this ventilator's optimal settings need to be changed as well, but it is usually not.

The physicians would like to see and adjust the settings of the mechanical lung ventilator often, multiple times a day, but this was not viable during the pandemic, as each such bedside visit required changing to a protective suit and back, for each patient separately. With the load of the hospitals, physicians did not have time for this and also needed to take care of other things, not just ventilation.

Long-term capacity overload leads to undesired psychological factors [5] – namely physician burnout. To help minimize the risk of cognitive overload and burnout, the proposed system's primary goal is to free up more of the physician's time to focus on other tasks needed to care for the patient by streamlining the checking of the ventilator screen.

The potential of such a system reaches far beyond its use during the pandemic. Mainly because it provides means for long-term analyzing long-term patient development and opens the possibilities of conciliar services when an opinion of a more experienced physician is required.

## 2 Previous work

We have developed a system [6,7], that allows a remote connection to the ventilator from a computer or phone of the physician. This connection streamlines the process because the physicians do not need to visit each patient preemptively. However, they can see what is happening in real-time from the clean physician's room (green zone) and only go to the contaminated (red zone) if necessary. An illustration of this system can be seen in figure 1.

The system works by first reading the HDMI/DVI signal from the ventilator output using a dedicated unit. This unit then makes a short 10s anonymized recording of the video signal and sends it to the server. This happens periodically, once each 15mins. Whenever a physician needs to see the output from a ventilator, he connects to the server, and using a simple web interface, he can find the latest recording or some older ones from the archive.

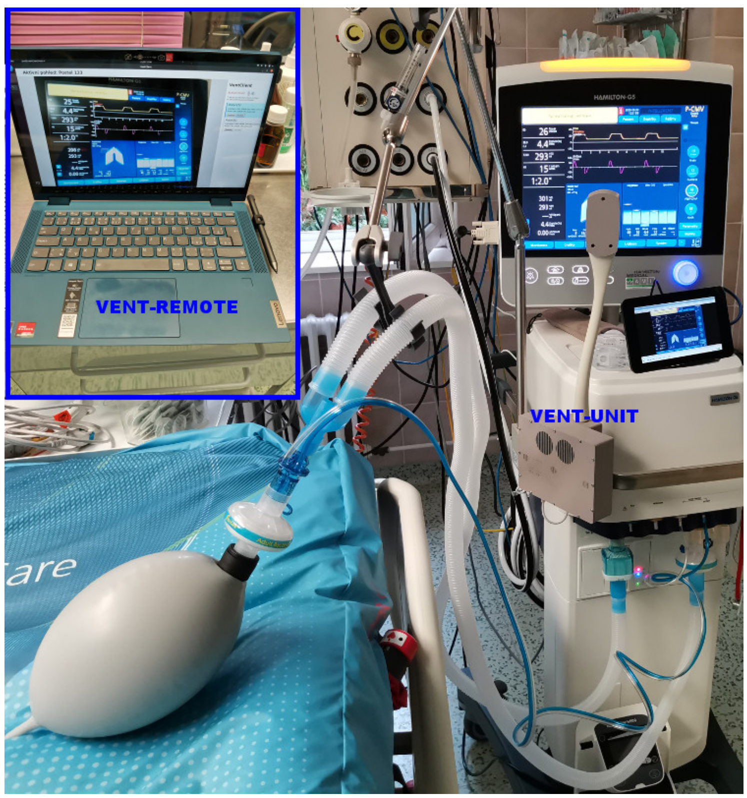



**Fig. 1.** First version of vent-unit system with camera module

There are already hospital information systems, for example, MetaVision [4], that collect and manage data from various sources, including bedside instruments. However, MetaVision is not able to connect to the ventilators directly and obtain the necessary data in the required detail. There is no option to connect the ventilator with MetaVision directly and provide live video or data stream. To the best of our knowledge, we are unaware of any other system that would provide this connection directly from the ventilator screen (or the screen of another hospital instrument.)

### 2.1 User story

As part of our previous work, we have conducted interviews with the procedure participants and we found that they had a very positive feedback – "I like the option of seeing the data remotely, including trends in time" or "It is like standing next to the ventilator, it's irreplaceable."

Nevertheless, despite positive feedback, there is still significant room for improvement, and we took this feedback to design and build a new system. Namely, the requested changes were

- The ability for live streaming
- Better web view
- An automated alert system with notifications about critical events
- Sharing one-time access to a monitor screen with another physician for the purpose of conciliar services.

## 3 The new system

From the ground up, we have designed a completely new system, addressing the requirements resulting from our previous work and an even stronger focus on security and reliability. The architecture is depicted in figure 2.

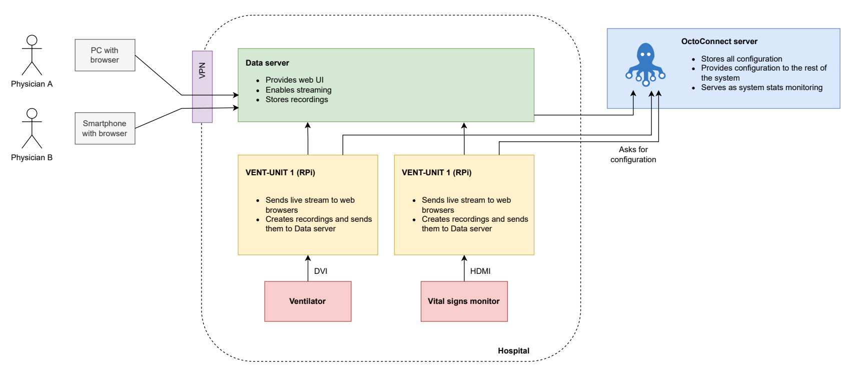



**Fig. 2.** The architecture of the new system

The system consists of three main parts: the device (VENT-UNIT), Data servers, and OctoConnect. Each connection between each part is mutually authenticated and encrypted. The system is designed such that each part can be temporarily offline, without a connection to the other parts of the system, and then recovers automatically.

We have built a completely new user interface – physicians have one point where to interact with our system, which is running our web app in a web browser. This web app now uses the SPA paradigm [2] and can be accessed both from mobile devices as well as the computers of the physicians. The devices (VENT-UNITs) can stream video data directly to this web app using WebRTC technology [3]. This stream does not require additional server bandwidth, as the stream is opened in a peer-to-peer fashion. The server is used only to help negotiate the connection between the device and the web app. However, the server can step in and relay the video stream when this direct connection fails (due to various NAT or firewall settings).

### 3.1 OctoConnect

OctoConnect (abbreviated as "OC") is our system that interconnects all devices and servers. All devices and data servers are connecting to the OC and asking it for configuration, software updates, and report statistics for the purpose of monitoring and timely response to any problem. OctoConnect is our administrative dashboard. It also serves as an authoritative database source for every connected part of the system (see figure 3). OC is also connected to our internal git repositories and automation and coordinates remote software updates.

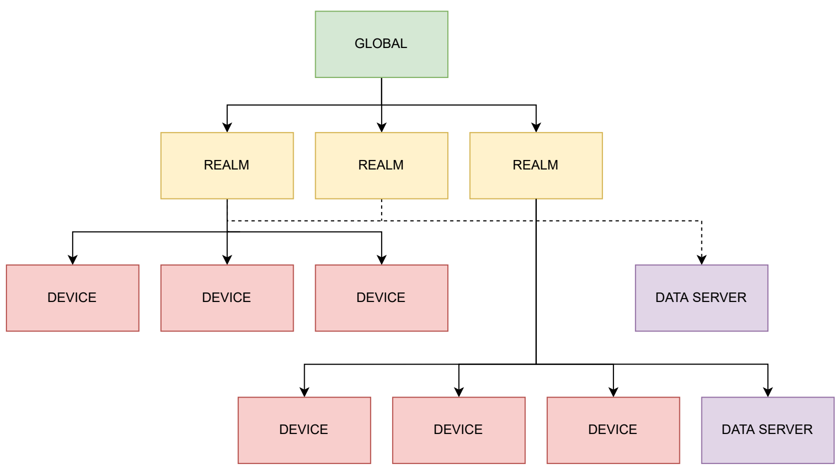



**Fig. 3.** An illustration of a hierarchical configuration

### 3.2 Software updates

We need the ability to update the software on the devices remotely; however, this can open the system to more possible cyber-security attacks or reliability issues. With this in mind, we had to be extremely careful about the design and implementation of this procedure.

The operating system in our VENT-UNITs is based on Buildroot. All of our software, as well as all other required packages, are compiled into one pre-built image. This image is then uploaded to the VENT-UNIT into the available slot from the A/B partition system [1]. The integrity and authenticity of the software image is verified, and then the system attempts to use it – by rebooting to this new slot. Suppose, however, this reboot fails for any reason or some part of the system is not functioning anymore (for example internet connection is failing). In that case, the image gets reverted, and the error is reported to the OC. The old system which worked before is still in charge.

## 4 Future work

The current problem with the ventilation we are facing is what we call dis-synchronies. When the artificial breathing from the machine interferes with the natural tendency of the patient to breathe. There are many types of these dis-synchronies, but what they have in common is that most of them are not directly detected by the ventilator, and they are not good in mid-to-long term (tolerable short term)

We want to develop a system that can detect an anomaly and notify the physicians so they can react to it and adjust the ventilation parameters accordingly.

We have more than 500 thousand (anonymized) recordings collected from multiple patients, beds, ventilator models, ventilation parameter settings, and other variations. We are currently working on machine learning models to classify any novel video stream with an appropriate alert level.

## 5 Acknowledgements

We would like to thank the members of the VentConnect team, including Lenka Vysloužilová, Jakub Vaňek, Jan Kubr, Milan Němý, and Eliška Niebauerová, for their contributions to this system.

We would also like to thank our colleagues at the University Hospital Královské Vinohrady, including Václav Zvoníček and František Duška, for their guidance and for letting us deploy the system in their hospital.

The research has been supported by institutional support of the Czech Technical University in Prague and University Hospital Královské Vinohrady. The project is funded by grant NU22-06-00625 – CheckMyScreen Project and partially by Research Center for Informatics (reg. no. CZ.02.1.01/0.0/0.0/16 019/0000765) supported by EU.

## References

1. Robust Auto-Update Controller – RAUC, `https://rauc.readthedocs.io/`

2. Single-page application – SPA, https://developer.mozilla.org/en-US/docs/Glossary/SPA
3. Web Real-Time Communication – WebRTC, https://developer.mozilla.org/en-US/docs/Web/API/WebRTC_API
4. iMDsoft: MetaVision ICU - clinical information system for critical care (2017), https://www.imd-soft.com/metavision-products/mv-for-intensive-care-adults-neonatal-and-pediatric
5. Iskander, M.: Burnout, cognitive overload, and metacognition in medicine. Medical Science Educator (2019)
6. Jakub Vaněk, M.M.: Ventconnect, system for remote monitoring of medical instruments. Proceedings of the 24th Bilateral Student Workshop CTU Prague. Dresden (2021)
7. Vysloužilová, L., Zvoníček, V., Duška, F., Jirman, J., Kubr, J., Lhotská, L., Macík, M., Němý, M., Niebauerová, E., Povišer, L., Samek, M., Vaněk, J.: Nová technologie v intenzivní péči pomáhá vzdáleně sledovat pacienty nejen s covid-19. Medsoft 33, 84–87 (09 2021)

# Nice noise: background noise enhancement with generated musical content

Pavel Husa[1], Šimon Prokop[2], Lukáš Kurilla[2]

[1]Faculty of Electrical Engineering, [2]Faculty of Architecture
Czech Technical University in Prague, Czech Republic

**Abstract.** In many environments, people are exposed to a certain level of background noise, e.g. in factories or public places. These conditions can cause irritation or distraction. We propose an interactive sound generator that will process noise and enhance it with harmonic and/or rhythmic content. In this phase, we discuss the possibilities for the generative music system and audio analysis system. Moreover, we present musical principles that the system can use, suggestions for evaluating the Nice Noise output, and ideas for future research.

## 1 Use case

A specific type of noise can be found in the industrial workplace environment, especially near automated production lines with conveyors and sinkholes, where the semi-finished product falls onto the surfaces of shutters. That usually produces semi-rhythmic impulses coupled with broader frequency bands of noise. Such qualities of the noise could be key in mitigating it by using a musical approach. In similar industrial workplaces, radio music is usually used to make long sessions of repetitive tasks more pleasant for the workers. However, finding consensus on the musical taste of employees can be difficult.

**Fig. 1.** Field recording of counting machines

## 2 Proposed design

The proposed design for the Nice noise system consists of two subsystems (see figure 2). The input analysis system will provide information about the noise in both the time and frequency domains. This analysis is crucial for identifying gaps in the noise in both domains and generating musical content to fill them. The first step in the process is to use active noise canceling to achieve the optimal noise-to-signal ratio to work with. In the initial phase, we plan to work with recordings from a factory. Additionally, we will use free datasets of noise environments, such as DagsHub [5] audio datasets or FreeSound [4], to test the proposed design on various datasets. The analysis system operates in two domains, frequency and time. We plan to perform a frequency analysis in the frequency domain to obtain information about the spectrum. This will help estimate the prevailing pitches in the signal that will be fed into the generation system. We will use beat detection algorithms on the time domain to detect if a rhythm is present. Subsequently, this information will be used to generate musical content. For the analysis task, we will test various audio analysis tools such as Essentia.js [**?**], which can be connected with machine learning models to yield better results, or Faust [3]. We will use MAX/MSP [2] for prototyping because it is well integrated with Ableton [1] and provides a good environment for tuning the analysis and music generation systems. The music generation system will be provided with information about the sonic spectrum, rhythm component, or irregular impulses if found. The musical principle for the rhythm will be call-and-response. If a significant impulse is detected, the system will "answer" to create a feeling of a steady rhythm. When an underlying steady beat is detected, the system will fill the missing gaps with a beat division to create a complete beat. Furthermore, when a strong fundamental frequency is detected, the system will react by creating an undertone, a "bass tone," to shift the context of the perceived sound. It is important to avoid monotonous output as employees listen to it for an extended period of time. This can be achieved by using multiple low-frequency oscillators (LFO) to modulate the output, e.g. the texture of the added tones.

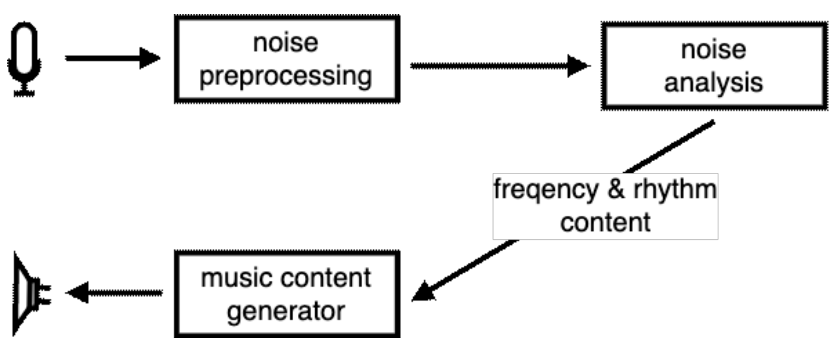



**Fig. 2.** Field recording of counting machines

## 3 Evaluation

Multiple evaluation approaches have been selected to assess the effectiveness of the generated musical content in enhancing the existing rhythmic and harmonic content in the noise. The first approach is subjective listening evaluation, which has been successfully employed in the evaluation of masking urban noises [7]. Additionally, noise annoyance score surveys will be conducted to gather data on the effectiveness of the generated musical content in reducing noise annoyance [6].

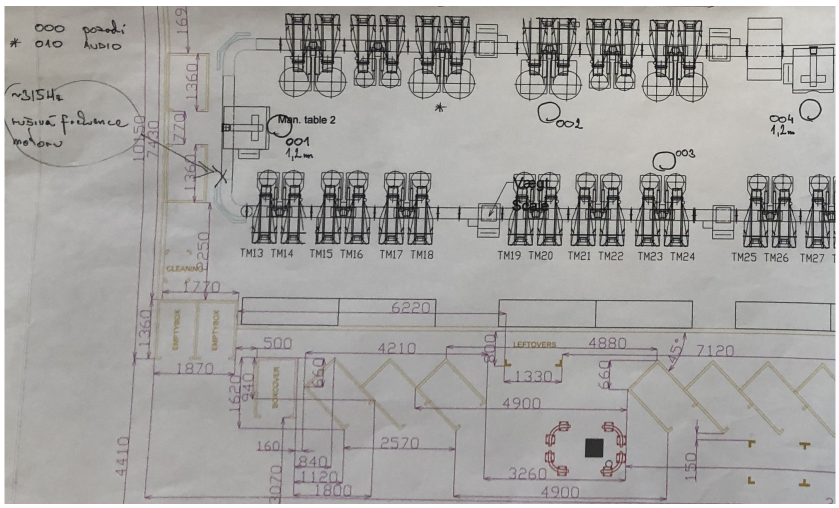



**Fig. 3.** Layout of measurement points of a double string of counting machines for small plastic pieces

Furthermore, music information retrieval (MIR) algorithms will be employed to provide useful information about the output. This approach considers the output sound as music, and MIR algorithms can yield comparable metrics that can be used between multiple iterations of our system or compared to the closest music genre, ambient music. Genetic algorithms will also improve these metrics since the system is fully parametric. This approach of combining different evaluation techniques, such as subjective listening evaluation, noise annoyance score surveys, and music information retrieval algorithms, will provide a thorough understanding of the effectiveness of the Nice noise project in enhancing the acoustic environment in industrial settings and potentially other public spaces.

## 4 Conclusion

The Nice noise system is an interactive sound generator that aims to enhance background noise in environments such as factories and public places by processing and adding harmonic and rhythmic content. The proposed design for the system includes an input analysis system and a music generation system, which will use active noise canceling and various audio analysis tools to analyze the noise and generate musical content to fill gaps in the noise. The effectiveness of the generated musical content will be evaluated through subjective listening evaluations and noise annoyance score surveys. This research presents a promising solution for mitigating background noise in industrial workplaces, and further research can be done to improve the system's functionality and effectiveness.

**Acknowledgments.** This work has been supported at CVUT in Prague by grant no. SGS22/075/OHK1/1T/15 Fakulta Architektury, Ústav modelového navrhování MOLAB and SGS22/172/OHK3/3T/13 (FIS 161 - 1612233C000), and by RCI (CZ.02.1.01/0.0/0.0/16 019/0000765)

## References

1. Creative tools for music makers | ableton, `https://www.ableton.com/`
2. Cycling '74, `https://cycling74.com/`
3. Faust programming language, `https://faust.grame.fr/`
4. Freesound - sound search, `https://freesound.org/search/?q=industrial+background+noise`
5. Open-source audio datasets, `https://github.com/DagsHub/audio-datasets`, original-date: 2021-10-05T15:06:49Z
6. Brink, M., Giorgis-Allemand, L., Schreckenberg, D., Evrard, A.S.: Pooling and comparing noise annoyance scores and "high annoyance" (HA) responses on the 5-point and 11-point scales: Principles and practical advice 18(14), `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8306719/`, publisher: Multidisciplinary Digital Publishing Institute (MDPI)
7. Shu, H., Song, Y., Zhou, H.: Assessment of music and water sounds for urban noise masking. In: TENCON 2018 - 2018 IEEE Region 10 Conference. pp. 1451–1455. ISSN: 2159-3450

# Parametric Curve Labeling

Václav Pavlovec

Faculty of Electrical Engineering,
Czech Technical University in Prague, Czech Republic

**Abstract.** The automatic label placement for line features is an important problem. In cartography, we find streets, rivers, or isolines in medicine veins or nerves, and graph edges in information visualization. In general, any thin-enough structure is treated as a line feature. Interactive scenarios require automated labeling of these features. However, existing methods do not consider several important aspects of the label layout, like visual interference with other visual features, proximity to other line features, or consistency of the layout style. Consequently, issues with legibility and label ambiguity can arise.I propose a method that aims to solve these issues. To provide immediate response in interactive scenarios, I first parametrize the line features and perform the criteria evaluation in the parameter domain.

## 1 Introduction

Visualization of textual attributes plays a vital part in various fields, such as Cartography, Information visualization, and Technical and Medical illustration, just to name a few. Three basic types, or classes, of visualized features, are generally distinguished – point, line, and area. The placement of textual annotation differs for each of these classes, but the problems tend to be NP-hard. If the features are small enough, we talk about point-feature labeling. The problem has received significant attention from the community, and many optimization strategies have been proposed. The same is true for the case when the features are areas; then, we talk about area-feature labeling. The last kind of feature is called line-features. These comprise any thin-enough structures that cannot be considered area-features. Unlike the first two problems, comparatively little attention has been given to this problem by the labeling community.

The Cartographic community has long-established guidelines for labeling line-features, which aim to improve the legibility and ambiguity of the placed annotations. For example, smooth sections of the line-features should be used. Further, the slope of the section should be minimal to increase legibility. In the case of labeling road networks, the annotations should not be placed over intersections. This way, we improve the legibility and possibly ambiguity of the labels. However, existing research is most concerned with optimizing the number of annotations, or exploring their usage in environments, such as interactive maps, 3D or XR, often neglecting to construct apt quality criteria for the annotation placement, as established guidelines are not always the most suitable for new environments.

Existing map services exhibit identical symptoms. The annotations are sometimes placed in unsuitable positions, excessively curved, and overlap other visual features, as examples in Figure 1 show.


(a)

(b) (c)

**Fig. 1.** (a) Curved and overlapping labels on Mapy.cz, (b) a curved label on Mapy.cz, (c) overlapping label on OpenStreetMap

## 2 The proposed method

The solution I propose is to consider the properties of the line-feature, as well as the properties of other features in the scene. The end goal is to increase the legibility and (un)ambiguity of the features. I adapt ambiguity criteria from previous work for this purpose [2].

The common way one can imagine the process of selecting the best annotation position for a given line-feature, is pushing the annotation along the line, as shown in Figure 2. At every step along the line, the properties of the annotation must be evaluated.

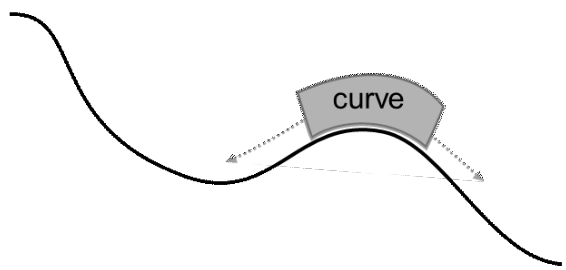


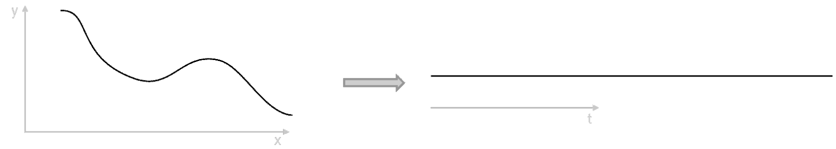

**Fig. 2.** Label sliding along a curve

The annotation must adhere not only to the established guidelines but also must reflect other features in the scene. Therefore, we need to evaluate the whole area the annotation would cover or affect. I propose a screen space method, which can be adapted for labeling a wide range of graphics. Consequently, every pixel of the annotation has to be taken into account.

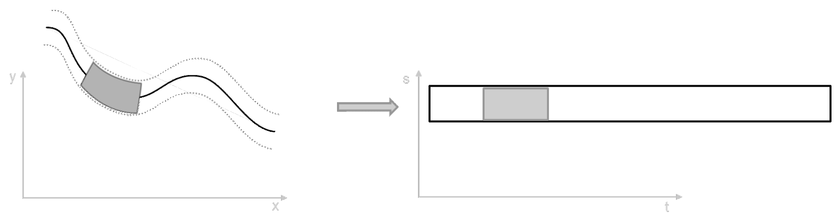
In previous work [2][1], we have dealt with this problem by employing Summed Area Tables (SATs), which allow us to evaluate all pixels of vertically or horizontally aligned annotations in constant time. However, in the case of line-features the annotations can have another alignment, as they are aligned to the line. I propose to transform the space around the line-feature into a rectangle, which would allow the use of SAT. See Figure 2 for an illustration of the process.

The criteria for label fitness can be evaluated in the transformed space of the parametrization and subsequently encoded as SAT. Since the evaluation of every pixel is independent of other pixels, the process can be massively parallelized by the use of GPUs. I list the possible criteria to influence label fitness below.

- Curvature
- Intersections
- Overlap of other visual features
- Label ambiguity
- Text orientation
- Zoom-consistency




(a)




(b)

**Fig. 3.** (a) Straightening of the line-feature in suitable parametrization, (b) space around the line-feature parametrized

## 3 Conclusion

I have proposed a method to evaluate labels for line features, which would take into account other visual features in the scene. This consequently leads to less ambiguous label positioning and eliminates overlaps between labels and other visual features. In order to speed up the evaluation, I propose a transformation of the image space to ”straighten” the line-features and allow the use of SATs, which speed up the selection of the best label candidates.

### Acknowledgement

This research has been supported by MEYS of Czechia OP VVV grant No. CZ.02.1.01/0.0/0.0/16_019/0000765 – Research Center for Informatics and by Grant Agency of CTU in Prague grant No. SGS22/173/OHK3/3T/13.

## References

1. Pavlovec, V., Cmolik, L.: Rapid labels: Point-feature labeling on gpu. IEEE Transactions on Visualization and Computer Graphics 28(1), 604–613 (2022)
2. Čmolík, L., Pavlovec, V., Wu, H.Y., Nöllenburg, M.: Mixed labeling: Integrating internal and external labels. IEEE Transactions on Visualization and Computer Graphics 28(4), 1848–1861 (2022)