Markus Seidl Thomas Moser
Kerstin Blumenstein Matthias Zeppelzauer
Michael Iber (Eds.)

FMT 2018

Proceedings of the 11th Forum Media Technology and 4th All Around Audio Symposium

November 28-29, 2018
St. Pölten, Austria

Editors' addresses:
St. Pölten University of Applied Sciences
Institute of Creative\Media/Technologies
Matthias Corvinus-Straße 15
3100 St. Pölten
Austria

{markus.seidl | thomas.moser | kerstin.blumenstein | matthias.zeppelzauer | michael.iber}@fhstp.ac.at

Preface

Audiovisual, interactive and mobile media interweave with our everyday life, a major aspect of the so-called megatrend digitalization. In almost all our areas of life, be it private or professional, they change our behavior and affect our perception, thinking and feeling. The 'Forum Media Technology (FMT)' is a scientific event that aims at providing an intensive dialogue of experts working at agencies, studios, and companies of different industries with students, lecturers, researchers, and developers in the field of digital media.

This year in its 11$^{\text{th}}$ edition, the Forum Media Technology was held on November 28-29, 2018 at St. Pölten University of Applied Sciences, Austria. In conjunction with the main conference, the 4$^{\text{th}}$ edition of the trans-disciplinary symposium 'All Around Audio (AAA)' were part of the two day program.

All Around Audio Symposium

Although, audio has not abandoned its status as a standalone discipline, its trans-disciplinary participation at the conception and design of products and environments has become more and more essential. In this sense, All Around Audio not only addresses specialists in the audio domain but particularly encourages researchers and designers from other fields to participate in the symposium. In its forth edition, 16 international speakers were invited for talks on a wide spectrum of topics reaching from media economy, audio technology, auditory display to music and media arts.

Forum Media Technology Conference Track

Submissions for the FMT conference track were accepted in two categories of full and short papers. All submitted papers underwent a double-blind review process, during which each paper was reviewed by at least three members of the international program committee (IPC) of the conference. Based on the written reviews, final decisions were made by the paper chairs and finally 9 papers (6 full + 2 short + 1 demo) were accepted for presentation at the conference (acceptance rate: 90%). The accepted papers can be clustered into three areas: audio, augmented reality & visualization, as well as human-computer interaction.

For the third time this year, also one best paper and two honorable mention awards were given to the top three contributions of the FMT. The award is based on the scientific excellence of the paper as well as the presentation performance of the authors at the conference. The selection process for the best paper award followed a three-step process. First, IPC members indicated whether a paper should be considered for the award. Based on that, the paper chairs decided on a short list of three nominated papers for the award and the selected papers' authors were informed about their nomination. The final decision was made by an anonymous best paper committee consisting of three members who took into account both the paper as well as the presentation to make their final decision.

November 2018

Markus Seidl
Thomas Moser
Kerstin Blumenstein
Matthias Zeppelzauer
Michael Iber

Best Paper Award Sponsor:

WISSENSCHAFT · FORSCHUNG NIEDERÖSTERREICH

Organizing Committee

Conference Chairs

Markus Seidl, St. Pölten UAS, AT
Thomas Moser, St. Pölten UAS, AT

Paper Chairs

Kerstin Blumenstein, St. Pölten UAS, AT
Matthias Zeppelzauer, St. Pölten UAS, AT

Graduate Consortium Chair

Grischa Schmiedl, St. Pölten UAS, AT

AAA Symposium Chair

Michael Iber, St. Pölten UAS, AT
Franziska Bruckner, St. Pölten UAS, AT

Local Organizers

Sylvia Petrovic-Majer, St. Pölten UAS, AT
Julia Klaus, St. Pölten UAS, AT

Program Committee

Christoph Anthes, University of Applied Sciences Upper Austria, AT

Thomas Artner, TU Wien, AT

Paulo Buono, University of Bari Aldo Moro, IT

Mario Döller, University Bremen, DE

Matthias Husinsky, St. Pölten University of Applied Sciences, AT

Michael Iber, St. Pölten University of Applied Sciences, AT

Andreas Jakl, St. Pölten University of Applied Sciences, AT

Patrick Lechner, St. Pölten University of Applied Sciences, AT

Munir Merdan, Practical Robotics Institute Austria, AT

Peter Plessas, Institute of Electronic Music and Acoustics Graz, AT

Alexander Schindler, AIT, AT

Sebastian Schlund, Vienna University of Technology, AT

Raphael Schneeberger, PocketScience GmbH, AT

Klaus Schöffmann, Klagenfurt University, AT

Katta Spiel, TU Wien, AT

Markus Taub, University of Applied Sciences Burgenland, AT

Markus Wagner, St. Pölten University of Applied Sciences, AT

Josef Wolfartsberger, University of Applied Sciences Upper Austria, AT

Michael Zeiller, University of Applied Sciences Burgenland, AT

Contents

Keynote

Keynote: Writer Identification on Historical Manuscripts

Robert Sablatnig, TU Wien, AT

Abstract

In recent years, Automatic Writer Identification (AWI) has received a lot of attention in the document analysis community. However, most research has been conducted on contemporary benchmark sets. These datasets typically do not contain any noise or artefacts caused by the conversion methodology. This article analyses how current state-of-the-art methods in writer identification perform on historical documents. In contrast to contemporary documents, historical data often contain artefacts such as holes, rips, or water stains which make reliable identification error-prone.

Biographie

Robert Sablatnig was born in Klagenfurt, Carinthia, Austria, in 1965. From 1992 to 2003 he was an assistant professor (Univ.Ass.), and from 2003 to 2010 an associate professor (ao Univ.Prof.) of computer vision at the Pattern Recognition and Image Processing Group. From 2005 to 2017 he was the head of the Institute of Computer Aided Automation. Since 2010 he is heading the Computer Vision Lab, which is part of the newly founded Institute of Visual Computing & Human-Centered Technology (TU Wien), engaged in research, project leading, and teaching. His research interests are 3D Computer Vision including Range Finder, Stereovision, Shape from X, Registration, Calibration, Robot Vision; Automatic Visual Inspection, Hierarchical Pattern Recognition, Video data analysis (Motion and Tracking), Automated Document Analysis, Multispectral Imaging, Virtual- and Augmented Reality, and Applications in Industry and Cultural Heritage Preservation.

Session 1: Audio

Exploration of Auditory Augmentation in an Interdisciplinary Prototyping Workshop

Katharina Groß-Vogt, Marian Weger, Robert Höldrich
Institute for Electronic Music and Acoustics (IEM)
University of Music and Performing Arts
Graz, Austria
Email: vogt@iem.at, weger@iem.at, hoeldrich@iem.at

***Abstract*—Auditory augmentation has been proposed as a specific, ambient sonification method. This paper describes an interdisciplinary workshop exploring this method by designing prototypes. Three of the prototypes are presented and discussed. Concluding on the workshop's results, the authors suggest a broader definition and deduce limitations of auditory augmentation.**

I. Introduction

Sonification, in the authors' definition, is the translation of information for auditory perception, the acoustic equivalent to data visualization [1]. Our institute pursues the Science by Ear (SBE) workshop series on sonification. Since 2005, four of these workshops took place. They explored sonification in an interdisciplinary, small group of attendees, and each had a different focus: The first workshop (SBE1) explored data sets in a variety of disciplines (from sociology to physics).[1] The workshop set-up has proven convincing, even if the large variety of disciplines and the scientific dialects of their communities were demanding for the attendees. SBE2 focused on physics' data[2] and SBE3 on climate data[3].

One of the lessons learned from hosting this series is to carefully balance the interdisciplinary, creative setting with well-prepared tasks. If this is achieved, the setting provides a fruitful development of prototypes, as shown in this paper. The layout of the workshop is discussed in Sec. II. In the fourth edition (SBE4), the focus was less on a specific data domain but instead on exploration of a specific sonification method: Auditory augmentation has been proposed by Bovermann et al. [2] as altering the characteristics of structure-borne sounds for the purpose of supporting tools for data representation. Besides this term, many notions and systems following a similar idea have been published, as discussed in Sec. III. As exemplary data set used in this exploration, we chose data of in-home electric power consumption. Section IV introduces our working definition of auditory augmentation and three of the prototypes developed on three hardware platforms. Finally, in Sec. V, we discuss some answers to our research questions, and conclude in Sec. VI.

[1]SBE1: http://sonenvir.at/workshop/
[2]SBE2: http://qcd-audio.at/sbe2.html
[3]SBE3: https://sysson.kug.ac.at/index.php?id=16451

II. Workshop layout

The workshops of the Science by Ear series are structured as follows. About 20 people with different backgrounds are invited to participate: sound and sonification experts, researchers of a certain domain science, and composers or sound artists. During the workshop, they take the roles of programmers, sound experts, data experts, and others (e.g., moderators of the teams). One workshop takes place on three or four days. After some introduction, participants are split into three to four teams of about five people. Each team is working on one sonification task with a given data set for three hours. Team members always include a moderator who also takes notes, and one or two dedicated programmers who implement the ideas of the team. The prototyping sessions combine brainstorming, data listening, verbal sketching, concept development, and programming. After each session, the teams gather in plenum to show and discuss their prototypic sonifications. Besides the hands-on character of the workshops, there is a certain challenge between teams to produce interesting results within three hours. Data sets, tasks, and software are prepared by the organizers in order to ensure that technical obstacles can be overcome within the limited time.

Within SBE4, the fourth edition of the workshop series, the method of auditory augmentation has been explored. This implied another level of complexity, as not only the data and the software needed to be prepared by the organizers and understood by the participants, but also possibilities and restrictions of the provided hardware systems had to be communicated.

Including the authors, 19 participants took part. Eleven of these can be counted to the sonification community (but with varying backgrounds in sciences, arts, and humanities), while the rest included two media and interaction experts, two composers, two sound engineers, one musicologist, and one sociologist. Participants divided in seven at pre-doc and twelve at post-doc level or above; in three female and 16 male participants. Not all of them took part throughout the whole workshop, leading to varying group sizes of three to six for the prototyping sessions.

III. Related work on auditory augmentation

The concept of auditory augmentation has been proposed by Bovermann et al. [2] as "building blocks supporting the design of data representation tools, which unobtrusively alter the auditory characteristics of structure-borne sounds.” One of these authors’ examples eliciting the concept is ‘WetterReim’ [3]. An ordinary laptop keyboard is equipped with a contact microphone, recording the typing sounds. The microphone signal is filtered with varying parameters that depend on the weather condition. The output is played back in real time and fuses with the original sound to a new auditory gestalt. In short: depending on the weather outside, typing on the keyboard sounds different.

The concept of auditory augmentation has been discussed to be extended to the more general *blended sonification* [4] which “describes the process of manipulating physical interaction sounds or environmental sounds in such a way that the resulting sound carries additional information of interest while the formed auditory gestalt is still perceived as coherent auditory event.” Blended sonifications “blend into the user’s environment without confronting users with any explicitly perceived technology”. In consequence, they provide an ambient sonification channel. Blended sonification is similar to or even encompassing auditory augmentation but takes into account environmental sounds, in addition to structure-borne sounds. A different generalization of auditory augmentation is given by Weger et al. [5] who “define *augmented auditory feedback* as the artificially modified sonic reaction to physical interaction”. Augmented auditory feedback can become an auditory augmentation if it conveys additional information. For the context of SBE4 we decided to stick to the original term *auditory augmentation* with a new working definition that incorporates the prepared platforms and tasks (see Sec. IV-A).

Looking at a broader context, auditory augmentation is part of *Sonic Interaction Design (SID)* which has been defined by various authors with different focuses: Rocchesso et al. [6] defined that it “explores ways in which sound can be used to convey information, meaning, aesthetic and emotional qualities in interactive contexts.” Franinović and Serafin [7] set the focus more on phenomenology and perception: “Sonic interaction design [...] [considers] sound as an active medium that can enable novel phenomenological and social experiences with and through interactive technology.” Auditory augmentation is certainly part of sonic interaction design, and it is within the scope of this paper to elicit the specificities about it.

We found a few more SID systems in the literature that are closely related to auditory augmentation, especially regarding its focus being an ambient display. For instance, Ferguson [8] developed a series of prototypes which are similar to the ones that emerged from our workshop. One example is a wardrobe door that plays back a sonification of the daily weather forecast when opened; Ferguson uses the term *ambient sonification systems*.

Kilander and Lönnqvist developed *Weakly Intrusive Ambient Soundscapes for Intuitive State Perception (WISP)* [9] in order to “convey an intuitive sense of any graspable process” rather than a direct information display. In a ubiquitous service environment, individual notifications are presented with a sound associated to the user. Playback volume and reverb are used to convey three levels of intensity of the notification. High intensity is mapped to a dry, loud sound, while low intensity is saturated by reverb, giving the impression of a far sound. Finally, a similar project has been realized by Brazil and Fernström [10] who explored a basic system for an ambient auditory display of a work group. The presence of colleagues is sonified as a soundscape of individual sounds, each time a person enters or leaves the workplace. The proposed system utilizes auditory icons [11, p. 325-338] for creating a soundscape and is not based on auditory feedback as was the case in WISP.

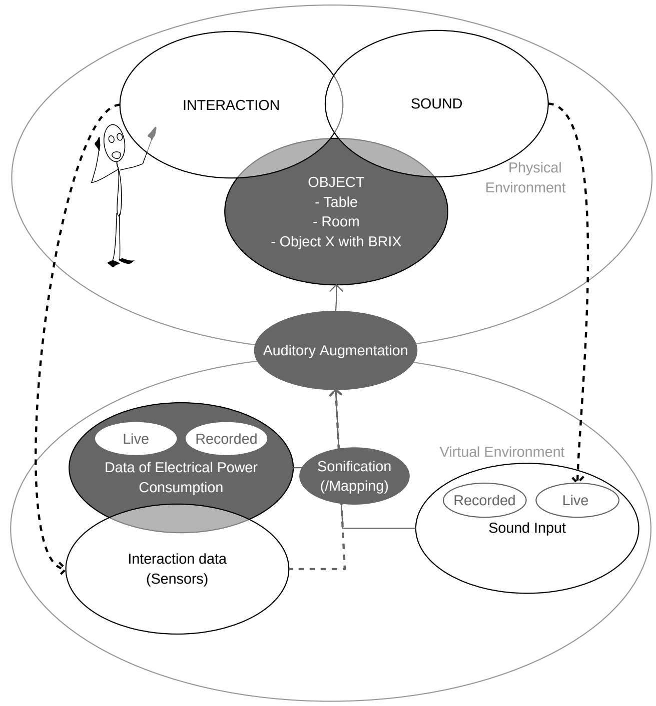


Fig. 1. Sketch of our working definition of auditory augmentation. Three mandatory elements are shaded in gray: an object delivering either sound input and/or interaction input, data, and sonification of these data which is auditorily augmenting the object in a closed feedback loop.

IV. Science by Ear 4

A. Working definition of auditory augmentation

SBE4 provided a set-up for exploring auditory augmentation and defined that

> Auditory augmentation is the augmentation of a physical object and/or its sound *by sound* which conveys additional information.

As sketched in Fig. 1, three elements are needed for auditory augmentation:

1) Physical *objects* in a physical environment. In our workshop setting, these were a table, a room, or any sensor-equipped physical object. These objects may produce a sound, or not; users might interact physically with them, or not. In some cases, the sound is a result of the interaction, but does not have to be. *Either* of these

inputs has to be there: real-time sound input *or* real-time data input from the interaction.
2) *Data* that are sonified. These can be real-time data or recorded data; in the setting of the workshop, we chose data sets of electric power consumption. The sonification may use 'natural' sound input (real-time sound or field recordings), or may be based on sound synthesis. Further input for parametrizing the sonification can stem from interaction data.
3) The *sonification* is played back in the physical environment, auditorily augmenting the physical object we started from.

In short, for auditory augmentation we need an object which produces sound or is being interacted with, data, and their sonification.

B. Interaction platforms

The various possibilities of auditory augmentation have been explored on three different platforms during the workshop.

1) ROOM: The ROOM is situated in the institute's main performance and lecture hall, equipped with a 24-channel loudspeaker array on a half-sphere for ambisonic sound spatialization [12]. Furthermore, there are five microphones mounted permanently to allow for a virtual room acoustics. For SBE4 we prepared to work with both live sound input from the microphones of the virtual acoustics system and additionally added ambient sounds.

2) TABLE: The TABLE is an experimentation platform developed within an ongoing PhD project (see [5]). Technically, it incorporates a wooden board or table (depending on its orientation in space) equipped with hidden contact microphones and exciters or additional loudspeakers; a marker-based optical tracking system locates the position of any object or hand interacting with the surface. Any sound produced on the TABLE is recorded and can be augmented in real time through a filter-based modal synthesis approach. The prepared setting for the workshop allows to change the perceived materiality in a plausible way while, e.g., knocking or writing on it.

3) BRIX: Our co-organizers (see Acknowledgment) provided their BRIX system [13], [14]. This system has been developed to allow for simple prototyping of interactive objects. In the prototyping sessions with the BRIX, the team could choose an interaction scenario with any object, equipping it with BRIX sensors and/or with microphones. Sensors that have been prepared for the workshop include accelerometer and gyroscope, as well as light, proximity, humidity, and temperature sensors.

C. Data sets

Next to the three hardware platforms described above, we prepared three data sets of electric power consumption. The data sets are of different nature (real-time vs. recorded data) and therefore propose different tasks, i.e., real-time monitoring as opposed to data exploration. The teams had to develop scenarios that support the saving of electric energy consumption.

1) REAL-TIME: The REAL-TIME data set was provided as a real-time power measurement of five appliances at our institute's kitchen (see [15] for how this system has been used before). Alternatively, the teams could attach the measurement plugs to any other appliances during the workshop. The sampling interval was about one second; data was sent over OSC, with a measurement range between zero and 3 000 Watt. Measured kitchen appliances were dish washer, coffee machine, water kettle, microwave, and fridge.

2) PLUGWISE: The PLUGWISE data set stems from a private household where nine appliances' loads have been measured for one year with a sampling interval of 10 minutes. Measured appliances comprise: kitchen water kettle, ceiling light and media station in the living room, fridge, toaster, dehumidifier, dishwasher, washing machine, and TV.

3) IRISH: The IRISH data set stems from a large survey of smart meter data in Ireland, collecting data of 12 000 Irish households over 1.5 years with a sampling interval of 30 minutes [16]. We extracted 54 households for each combination of three family structures (single, couple, family), two education levels, and two housing types (apartment vs. detached house).

D. Resulting prototypes

During the three days of the workshop, four parallel prototyping sessions took place; one for each of the three data sets and an open session on the last day where chosen prototypes were refined. An overview of all the resulting prototypes is shown on the SBE4 website[4]. In this paper, we only focus on three exemplary cases: *sonic floor plan*, *writing resonances*, and *standby door*, realized with the three different platforms respectively (see Fig. 2).

1) Sonic floor plan (ROOM): The real-time data set used in this scenario provided data of electric power consumption of different devices in one specific household. The team developed a scenario with an assumed floor plan of the household (see Fig. 2a). Feedback on energy consumption is played back in one room (e.g., a media room) on a surround sound system. A sound occurs periodically after a specified time, as well as when a person enters the room. The appliance that currently consumes the most energy defines the direction for sound spatialization. Environmental sounds from outside the building are captured by a microphone. These are played back in the room with loudness and position depending on the level of energy consumption. As only the power consumption of the appliance with the highest load is sonified, even small standby consumption may attract attention, e.g., when no major energy consumer is active.

2) Writing resonances (TABLE): Writing resonances is the most 'classical' example of auditory augmentation during the workshop, because it is based on structure-borne sounds and therefore clearly fulfills the initial definition of Bovermann et al. The scenario is to provide feedback of in-home power consumption through an auditorily augmented writing desk (see Fig. 2b), based on the system presented in [5]. The

[4]SBE4: https://iem.at/sbe4

table is augmented through additional resonances, based on a physical model. The size of the modeled plate is controlled in real time by the total amount of electric power consumption, employing the metaphor of a larger table (i.e., deeper sound) when more power is consumed. With the augmentation being only a modulation of the existing interaction sounds, the sonification only gets audible through interaction with the table; the feedback is therefore calm and unobtrusive. The primary task is writing, but also an active request of information by knocking is possible. This prototype has been extended in the fourth, open session, allowing for a modulation of individual partials in order to additionally convey information concerning the different appliances' individual power consumption.

3) Standby door (BRIX): This scenario augmented our institute's inside entry door (see Fig. 2c for the prototypic mockup). Most potential for saving energy lies in the reduction of standby consumption. Therefore, when the door is opened, e.g., in the morning, the standby consumption over the recent past is communicated through a simple parameter mapping sonification. The playback speed depends on the assumed stress level of the user, derived from the opening speed of the door. In the open prototyping session, this approach has been extended to be able to sonify individual appliances. For intuitive discrimination between them, the sonification is based on synthesized speech. When opening the door, an emotionless computer-voice says 'coffee machine', 'microwave', and the like, with timbre, loudness, or other parameters controlled by how much electricity this specific appliance has used over night.

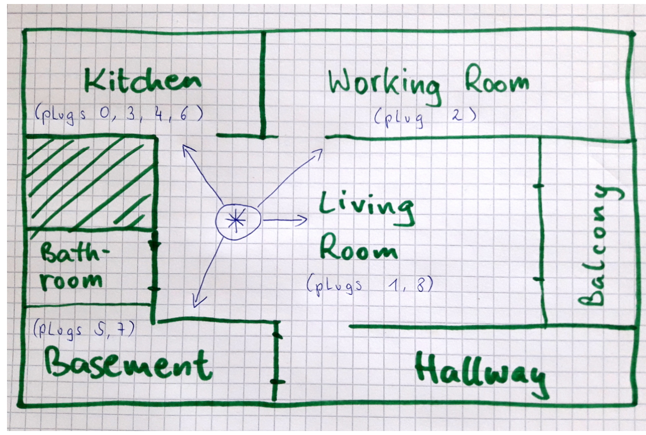
V. Feedback and Analysis

We explored four main research questions by applying various analysis methods on the results of the workshop. Plenary sessions with discussions of the presented prototypes have been recorded and partially excerpted. Additional inputs are written notes, code, and demo videos resulting from the prototyping sessions. All these inputs led to fundamental considerations on auditory augmentation and how it can be used, but also to general feedback concerning sonification, workshop setting, or design issues.

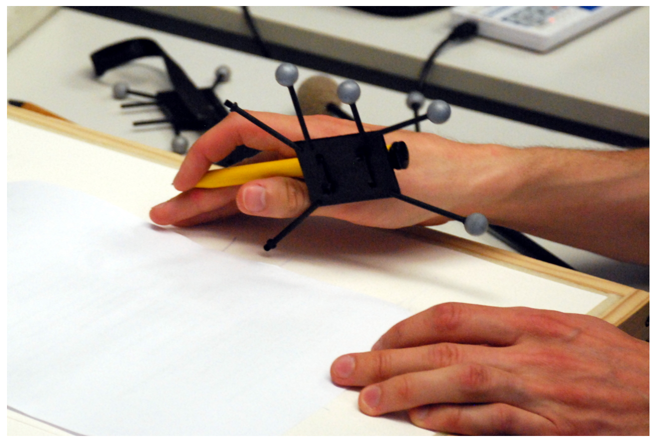
A. Peculiarity of sound in augmented reality

Augmented reality usually refers to a mainly visual system. But if this concept is transferred to audio, why then is listening to the radio not 'auditorily augmented reality'? Or is it? The underlying question is: what are the peculiarities of sound in the context of augmented reality?

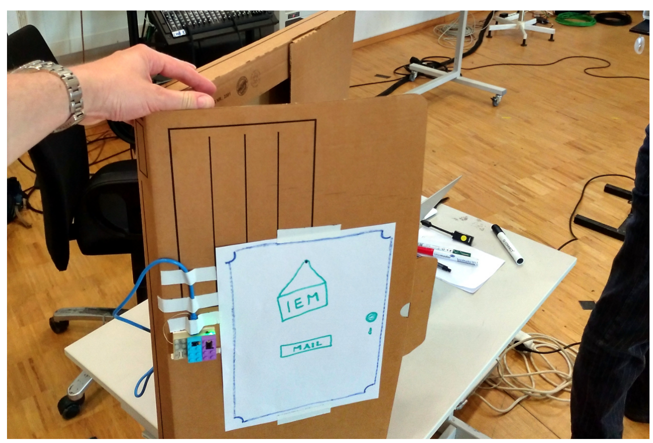
Concerning the radio, the answer is relatively clear. In the visual domain, augmented reality usually does not include an overlay of a video on top of the visually perceived scene if there is no direct connection between them [17]. Applying this argumentation to the auditory domain, the radio (being an overlay to the acoustic environment) does not directly interact with the physical environment of the user and therefore can not be regarded as augmented reality.



(a) *Sonic floor plan.*



(b) *Writing resonances.*



(c) Standby door.

Fig. 2. Three of the prototypes that have been developed during the workshop.

A more profound analysis of this question is without the scope of this paper. Still, a few things came up during the discussions in the workshop. On the one hand, sonification is challenging because visual and auditory memory work differently, and therefore two sonifications are more difficult to compare than two static visualizations. These and other challenges are well known, see for instance [18]. On the other hand, designing for sound creates new perspectives, e.g., on the quality of an interaction. One participant reported, his group had discussed the nature of an interaction in their scenario (in the *standby door* prototype, the quickness of opening the door is related to the user's mood). Even if sound is not involved directly, thinking about the design is different with sound "deep in our minds" (participant P10).

To conclude, 'auditorily augmented reality' clearly behaves different from its visual counterpart, and this fact deserves more systematic research.

B. Definition and limitations of auditory augmentation

One purpose of the workshop was to develop and test our working definition which is deliberately wide to incorporate different platforms. The question is, if this definition is more useful.

Our working definition of auditory augmentation has not been questioned within the workshop; therefore, we would like to propose it for future applications. Still, one aspect deserves more attention than has been discussed above, as came up during the final discussion round. The definition starts from an object that is being interacted with, i.e., a primary task of the user with the object is pre-assumed. "Having a concrete task helps to design" (P1), and helped the teams to elaborate their scenarios during the sessions. During the final feedback round, however, it has turned out that *the task* is an ill-defined entity. Does it relate only to the interaction between user and object? And if the object is augmented, and its sound conveys additional information, is there such a thing as a monitoring task? Does only a goal-oriented activity comprise a task or can it be a by-product of daily "state of being" (P7)?

We conclude that auditory augmentation always involves a primary task, may it be goal-oriented or not (e.g., writing on a desk in the *writing resonances* prototype or just being in the media room of the *sonic floor plan*). This task should not be disturbed by the augmentation, but rather the augmentation adds a secondary task of monitoring.

C. Relationship between sound, data, and augmented object

In addition, we aimed at exploring the qualitative factors between object and sound in the context of auditory augmentation. Which qualities are needed to (fully) describe their relationship? For example:

- Is the sound, in relation to the object, plausible? Is the mapping of data to sound intuitive/metaphoric?
- Is the augmented object more useful, or more fun than the original one? Does its affordance change and is the original interaction, i.e. the primary task of the user with the object, disturbed?

A central issue that has been raised throughout the workshop was in how far objects change in perception when they are auditorily augmented. This experience, "you can only get it if you interact yourself" (P1).

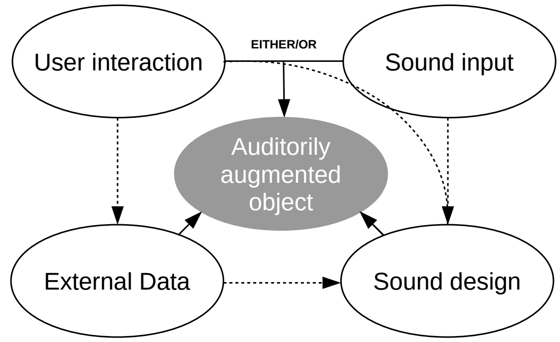


Fig. 3. An auditorily augmented object is influenced by three elements: either of the user interaction or the object's sound, external data, and the sound design used in the sonification.

For comparison between the three exemplary prototypes (Sec. IV), we analyzed the qualitative behavior of input elements on auditory augmentation. Abstracting from Fig. 1, we identify four of these elements: user interaction, input sound, external data, and sound design. This more abstract concept of auditory augmentation is shown in Fig. 3. We assume that more coherent relations between user interaction, input sound, external data, and sound design lead to higher naturalness and intuitiveness of the auditory augmentation system. Multiple dependencies are possible, even though not all are needed for auditory augmentation:

- User interaction may influence external data, e.g., turning on the water kettle while its electric power consumption serves as data. The task of writing by hand, however, does not directly modulate in-home electric power consumption.
- Data may have a close link to sound design, utilizing either direct sonification (e.g., audification [11, p. 301-324]) or a fitting metaphor, e.g., a larger desk with lower pitch representing a larger energy consumption.
- User interaction may directly influence sound design, in all cases where sounds are augmented that have been produced by the interaction.
- Sound input may not be stemming from the interaction but from an external source; still it may influence sound design, e.g., when using environmental sounds from outside the windows as in the *sonic floor plan* prototype.

In conclusion, it seems that a more natural prototype of auditory augmentation has more coherent relationships between user interaction and/or sound input on the one hand, and external data and sound design on the other hand.

D. Perceptual factors of data and sound

As a final research question, we aimed at sketching out perceptual factors of data and sound concerning auditory augmentation systems. What is the capacity of information that can be conveyed with auditory augmentation? Which data are suitable for it? Which factors play a role for blending the augmented object in the environment: is it unobtrusive but salient enough in order to be perceived? Of course, a full analysis of perceptual factors can only be a result of evaluation that was not within the scope of the workshop. However, some ideas emerging from the final discussions may serve as a basis for future investigations.

As one participant articulated, auditory augmentation works best with low-dimensional data – "otherwise you are not augmenting the object but creating a nice sonification" (P7). The information capacity, i.e., the level of information that can be conveyed, is rather low for the examined platforms. Especially for ROOM and TABLE, only the reverberation of the room – respectively the resonances of the structure-borne sounds – can be changed, with only few levels that can be differentiated perceptually. The BRIX system is more flexible in design, therefore no general conclusion can be drawn. Sonifying is difficult under these conditions, because "it boils down to the question which dimensions you chose and which ones you leave out" (P11).

In the *writing resonances* prototype, the developing team found a borderline for perception. Depending on the quickness of parameter changes, sounds lost their gestalt identity with the interaction, i.e., sometimes the sounds were perceived as separate auditory events played from the loudspeakers – despite measured round-trip latency below 5 ms. This example shows that perception is very sensitive and systems need to be well evaluated.

E. General feedback

Next to finding some insights on the aforementioned research questions, the analysis of SBE4 provided some feedback on the workshop setting itself, as well as on general design issues.

A general issue of sonification is its right to exist – referring to Supper's thesis "about community's struggles to have listening to scientific data accepted as a scientific approach" [19]. Useful and meaningful sonifications are difficult to come up with, which raised the question, "why sonify at all"? It was part of the workshop design to develop useful scenarios for the pre-defined platforms. This worked well for the prototypes presented in this paper, but not for all.

Generally, feedback on the workshop design was positive. The interdisciplinary, hands-on sessions were "so enriching" (P1). Participants further reported that they had "really time to try out something" (P2) – even if a certain approach did not work out in the end, there was learning even by dead ends. However, it was remarked that the prepared platforms had "narrowed down" (P3) possible paths of design. Some participants articulated the wish to be more free in designing interaction scenarios independently from a platform (P3, P5), but the prepared setting was very time-efficient. Most prototypes have reached a promising state after the three hours' sessions: "there are nine prototypes that are really worthwhile considering and working on in the future" (P4).

One participant (P1) raised the issue that designing ambient displays means designing for the background, while in the designer's mind the sound is in the foreground: "we have an excitement for sound and sonic display". Therefore, it would be important to cultivate a beginner's mind – something, he stated, that has been well achieved with the interdisciplinary workshop setting for most of the participants. Another general design-issue concerns the difference between prototypes and final products. The realized prototypes are meant to be listened to for a longer period of time, but designers only hear them for a short period of time (P8). Furthermore, for the purpose of demonstration and presentation, prototypes need to exaggerate, while in final products the appropriate settings are usually less salient. Participants who worked with iconic sounds of audio recordings stated that some cartoonification is needed (e.g., through post-processing or re-synthesis of the sounds), because "for ordinary people they all sound the same" (P9).

VI. Conclusion

Within this paper, we conclude on results from an interdisciplinary workshop exploring the concept of auditory augmentation. The workshop resulted in nine prototypes and, among others, recorded discussions that have been analyzed. Concluding on this material, we propose to use the term auditory augmentation with a new, broader definition: auditory augmentation is the augmentation of a physical object and/or its sound *by sound* which conveys additional information.

General considerations for auditory augmentation are summarized as follows.

Auditory augmentation requires a primary task of a user with an object; this task is not explorative data analysis. One reason is that data for auditory augmentation needs to be low-dimensional. Another reason is the differentiation between auditory augmentation and sonification. By augmenting an object auditorily, a secondary task of monitoring in the background turns up. This task must not interfere with the primary task.

There seems to be a quality of 'naturalness' (affecting also the 'intuitivity') of systems of auditory augmentation. The most natural systems have several coherent relationships between the four possible input factors, user interaction and/or sound input, with external data and sound design. We envisage exploring this hypothesis further.

There are borderline cases of perception, where the fusion of auditory gestalts between the original sound and the augmented one does not work anymore. The influencing factors need to be explored systematically.

Finally, the analysis of the final discussions during the workshop proved that the developed workshop setting is convincing. It establishes an interdisciplinary, playful atmosphere of research by design. The balance of possible ingenuity and well-prepared tasks, platforms, and data sets are crucial for a successful event.

Acknowledgment

We would like to thank our co-organizers from CITEC, Bielefeld University, and all participants of SBE4:
Lorena Aldana Blanco, Luc Döbereiner, Josef Gründler, Thomas Hermann, Oliver Hödl, Doon MacDonald, Norberto Naal, Andreas Pirchner, David Pirrò, Brian Joseph Questa, Stefan Reichmann, Martin Rumori, Tony Stockman, Leopoldo Vargas, Paul Vickers, and Jiajun Yang.

References

[1] K. Vogt, "Sonification of simulations in computational physics," Ph.D. dissertation, Graz University, 2010.

[2] T. Bovermann, R. Tünnermann, and T. Hermann, "Auditory Augmentation," *International Journal on Ambient Computing and Intelligence (IJACI)*, vol. 2, no. 2, pp. 27–41, 2010.

[3] T. Bovermann. (2010) Auditory augmentation – wetterreim. [Online]. Available: https://vimeo.com/19581079

[4] R. Tünnermann, J. Hammerschmidt, and T. Hermann, "Blended sonification – sonification for casual information interaction," in *Proceedings of the International Conference on Auditory Display (ICAD)*, 7 2013.

[5] M. Weger, T. Hermann, and R. Höldrich, "Plausible auditory augmentation of physical interaction," in *Proceedings of the International Conference on Auditory Display (ICAD)*, 2018.

[6] D. Rocchesso, *Explorations in Sonic Interaction Design*. COST Office and Logos Verlag Berlin GmbH, 2011.

[7] K. Franinovic and S. Serafin, *Sonic Interaction Design*, K. Franinovic and S. Serafin, Eds. MIT Press, 2013.

[8] S. Ferguson, "Sonifying every day: Activating everyday interactions for ambient sonification systems," in *Proceedings of the International Conference on Auditory Display (ICAD)*, 2013.

[9] F. Kilander and P. Lönnqvist, "A whisper in the woods – an ambient soundscape for peripheral awareness of remote processes," in *Proceedings of the International Conference on Auditory Display (ICAD)*, 2002.

[10] E. Brazil and M. Fernström, "Investigating ambient auditory information systems," in *Proceedings of the International Conference on Auditory Display (ICAD)*, 2007.

[11] T. Hermann, A. Hunt, and J. G. Neuhoff, *The sonification handbook*. Logos Verlag Berlin, Germany, 2011.

[12] J. Zmolnig, W. Ritsch, and A. Sontacchi, "The iem-cube." International Conference on Auditory Display (ICAD), 2003.

[13] S. Zehe, "Brix$_2$ – the xtensible physical computing platform," https://www.techfak.uni-bielefeld.de/ags/ami/brix2/, 2014.

[14] ——, "BRIX$_2$ – A Versatile Toolkit for Rapid Prototyping and Education in Ubiquitous Computing," Ph.D. dissertation, Bielefeld University, 2018.

[15] K. Groß-Vogt, M. Weger, R. Höldrich, T. Hermann, T. Bovermann, and S. Reichmann, "Augmentation of an institute's kitchen: An ambient auditory display of electric power consumption," in *International Conference on Auditory Display (ICAD)*, 2018.

[16] Commission for Energy Regulation (CER). (2012) CER smart metering project – electricity customer behaviour trial, 2009-2010 [dataset]. Irish Social Science Data Archive. SN: 0012-00. [Online]. Available: www.ucd.ie/issda/CER-electricity

[17] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoperators & Virtual Environments*, vol. 6, no. 4, pp. 355–385, 1997.

[18] G. Kramer, B. Walker, T. Bonebright, P. Cook, J. Flowers, N. Miner, J. Neuhoff, R. Bargar, S. Barrass, J. Berger *et al.*, "The sonification report: Status of the field and research agenda. report prepared for the national science foundation by members of the international community for auditory display," *International Community for Auditory Display (ICAD), Santa Fe, NM*, 1999.

[19] A. Supper, "The search for the "killer application": Drawing the boundaries around the sonification of scientific data." in *The Oxford Handbook of Sound Studies.*, K. E. Pinch, T. & Bijsterveld, Ed. Oxford University Press, 2011.

cosy:sonics - A Mobile App to Explore Technology Reflection Among Students

Oliver Hödl[1], Peter Oberhauser[1], Peter Reichl[1,2]
[1]Cooperative Systems Research Group, Faculty of Computer Science, University of Vienna, Austria
[2]VRVis Center for Virtual Reality and Visualization Vienna, Austria
oliver.hoedl@univie.ac.at, mail@peteroberhauser.at, peter.reichl@univie.ac.at

***Abstract*—This paper describes the design, implementation and evaluation of the mobile application *cosy:sonics* to explore technology reflection. To this end, three different approaches have been combined: (1) a cryptocurrency parody called *EatCoin* aiming at a direct experience of the useless resource consumption caused by Bitcoin's Proof-of-Work concept, (2) an ultrasonic communication channel circumventing the protection barriers normally provided by a smartphone's flight mode, and (3) a Turing test variant demonstrating the current state of AI-based music composition. The resulting prototype has been tested during a live demo in a scenario triggering intuitive user interest in using the app and has additionally been reviewed by a group of experts from different domains. While it turns out that the application successfully prompted discussion and reflection among the participants, also some weaknesses could be identified during the trials that are considered in future work.**

I. Introduction

While the so-called "Digital Revolution" is rapidly progressing, critical reflection of the related upcoming technologies is becoming more and more important. Consider, for instance, the Cambridge Analytica scandal which has led to personal data of millions of Facebook users being harvested to analyse political views and use this intelligence for political elections and votes [1]. Contrary to all expectations, people nevertheless still "appear to relate well to Facebook both as a technology (in terms of functionality, reliability, and helpfulness beliefs) and as a 'person' (in terms of competence, integrity, and benevolence beliefs)"[2], despite of the fact that experts recommend to simply delete all social media accounts right away [3].

To explore ways for further triggering technology reflection, we focus on two principles: Firstly, we believe that, for the purpose of such discussions, potential dangers and consequences, which are inherent to future technology, need to be made as tangible as possible and strictly driven by direct user experience. Secondly, we believe that emphasis needs to be made to trained experts, i.e. mainly computer scientists themselves, as those are the people which eventually are in a position to change technology in a way that is more compatible with a human-centric perspective on the world of tomorrow.

Thus, we study the phenomenon of critical technology awareness among students using a mobile app that combines three recent technologies in a carefully crafted joint way. We address the application level, the content level and the level of underlying communication technology at the same time: First of all, our app named *cosy:sonics* closely resembles a cryptocurrency like Bitcoin, and targets especially its Proof-of-Work concept. However, instead of consuming enormous amounts of electrical energy for basically no purpose like Bitcoin does [4], our app consumes time, more specifically, the precious time of people attending a social get together event. To do so, we have implemented a puzzle which forces users to compare different pieces of classical music and decide whether each of these pieces has been created by a human composer or by an AI-based algorithm. Finally, we use the concept of ultrasonic communication, which is some sort of near field communication, that uses inaudible sound to transfer data. In contrast to Bitcoin, we anticipate that many people have not heard about this technology and are not aware of the privacy issues they are confronted with when using smartphones, for instance [5]. We chose the concept of Bitcoin mining and ultrasound communication deliberately to explore two different examples and different views on a critical reflection of technology.

To stimulate an intuitive incentive for using the app, we have chosen a get together event as basic scenario where participants are required to solve the mentioned AI puzzles to get some food. While in the original scenario, using the app and successfully solving the puzzle has been a requirement to attend the event at all, in our trial successful users have been awarded a so-called *EatCoin*, i.e. some piece of chocolate. The target group to study critical technology awareness are mainly computer science students. We expect them to have an above-average interest and knowledge concerning new technologies. Furthermore, we consider them the future experts within the domain. The study itself was embedded at a faculty event where students could participate voluntarily by using an app. Additionally, the app prototype was presented and discussed in a focus group with experts from the Austrian Federal Ministry for Transport, Technology and Innovation, considered as the actual experts who advise the government and work for political decision makers in relation to new technologies.

The remainder of this paper is structured as follows: After summarising the State of the Art in Section II, we describe our design and development methodology in Section III. Section IV presents the final prototype, while Section V is devoted to the user trials. We discuss results in Section VI, and finally Section VII provides our conclusions and possibe further work.

II. State of the Art

Our work encompasses issues from various different fields. We start to examine studies in the field of technology awareness and critical reflection. Then we review related work in the area of ultrasonic communication and Bitcoin. Both fields have potential issues that concern the awareness and critical reflection of these technologies.

A. Technology Awareness and Critical Reflection

Reflection is a process of learning, that ideally leads to a deeper understanding of an issue and more complex knowledge about it. Kori et al. [6] argue that technology can be used to enhance the effectiveness of reflection processes while Slovak et al. [7] underline the importance of design for the process of reflection.

Lin et al. [8] identified four characteristics of design features in technology that can help supporting student reflection: (1) *process displays* which explicitly show and explain issues and procedures to learners. (2) *process prompts* which prompt students to develop, explain and evaluate their own solutions. This could be achieved by posing the right questions and leading the learners towards answers. (3) *process models* which provide support to the learner in establishing own knowledge by providing model processes of how an expert would approach solving a similar problem. (4) *forum for reflective social discourse* which incentivises reflection as a social activity among students. The main goal is to trigger meaningful discussions that lead to new insights among the participants.

How people can be mislead or how long they are not aware of a certain dangerous technology, has been demonstrated with the Cambridge Analytica scandal [1]. During Facebook's early days of growth and gaining popularity in 2008, most people trusted the platform blindly [9]. Ten years later, in 2018, the Cambridge Analytica scandal initiated an exodus of users deleting their Facebook accounts but the majority of the users still remain on the platform [10].

B. Ultrasonic Communication

An example of a powerful new technology many people have not even heard about, is ultrasonic communication. Deshotels [11] analysed the general suitability of inaudible sound as communication channel to transmit data. Humans with ideal hearing can hear frequencies approximately ranging from 20 Hz to 20 kHz. The range reduces noticeably with the age of the person. Smartphones can produce and detect a significant amount of (for humans) inaudible sounds that can potentially be used for the transmission of information. However, a precise range of by smartphones detectable sounds cannot be determined due to the lack of standards among manufacturers [12].

Due to the unorthodox nature of transmitting data acoustically, safety precautions implemented in most operating systems do not target inauble communication. Deshotels [11] experimented in his study with methods to bypass security mechanisms on an Android smartphone. In his scenario he wanted to send sensitive data from a software-protected phone to a receiver. He secured the phone with standard Android measures and various versions of third-party software. However, he found out that most tools do not prevent sensitive data being exfiltrated acoustically, i.e. via inaudible sound, because access to the phone's speaker is usually not considered dangerous or privacy sensitive. There is no way to regulate permissions on which applications are allowed or prohibited of using the phone's speakers.

In a study published in 2017, the authors revealed that a small number of retail stores in two European cities already use ultrasonic beacons, as they call it, for location tracking [13]. Mavroudis et al. [14] studied the privacy and security of such ultrasound ecosystems. They critique the lack of security and privacy measures in mobile operating systems which allow severe "proximity marketing and advanced cross-device tracking techniques" mostly unknown and not recognisable by users.

Among the early approaches to counter these privacy and security issues with inaudible sound communication are Zeppelzauer et al. [5]. They implemented and presented, *SoniControl*, the first ultrasonic firewall app to detect different kinds of ultrasonic signals in real-time.

Two interactive music performances, *Sense of Space* [15] and *Poème Numérique* [16], use ultrasound communication to realise audience participation with smartphones. Both studies, however, focus on audience participation itself, the audience's experience and technological issues with ultrasound. They do not study or discuss privacy and security issues. *Poème Numérique* uses an open source app developed by Bartmann [17]. His work and the app itself provide detailed information and insight which is publicly available on how inaudible communication technology is realised.

Apart from security and privacy issues, it is not entirely clear how inaudible sound affects the human body. However, it is reported that, especially the lower spectrum of high frequency sound (that can, at least partially, be heard by humans) causes symptoms like annoyance, headache, tinnitus, fatigue and nausea. Moreover, ultrasonic signals with high sound pressure may cause hearing damage [18]. These adverse effects were observed only when people were exposed to high frequency noise at the workplace continuously, over a longer period of time. Exposing subjects to low pressure sound only for a short period of time should therefore be harmless.

C. Cryptocurrencies

In contrast to ultrasonic communication, cryptocurrencies are publicly well-known. Their gaining popularity, and in particular those of Bitcoin, and their developments according to market prices initiated a broad media coverage recently.

Our focus is not Bitcoin or any other cryptocurrency itself, but the fact, that Bitcoin mining causes an increasing energy consumption. The energy needed worldwide for the underlying Proof-of-Work concept is nearly as high as the power needed of the country of Denmark [19]. This seems to be problematic in two ways. First, the Bitcoin mining process, which uses blockchain technology to solve complex mathematic equations, is more or less useless for anything else than the Proof-of-Work awarded with Bitcoins [20]. Second comes the problem with the increasing energy consumption which is aggravated by the fact that primarily carbon is used to power Bitcoin mining. This is a contradiction to the global commitment to mitigate greenhouse gas emissions [21].

III. Design Aspects and Development

We used the open source app *Poème Numérique* as basis for our prototype app [17]. As the original app realised an avant-garde music performance with audience participation [16], we only use the basic app structure and the implementation of the ultrasonic communication technology, but completely re-designed the interaction design and the user interface for our study purposes. Contentwise we adapted the audience participation concept as well, but not for a music performance and with a complete different content.

The focus of our prototype was to develop an aesthetically pleasing user interface that is intuitively to use and provides enough information to explain the underlying ultrasonic communication technology. However, at the same time we try to keep it as simple and plain as possible to avoid any unnecessary content.

A. User Interface Design

For the design of the prototype, we followed the fundamental design principles of Norman [22]. In the following, we describe how our app design incorporates these guidelines. In general, every view of the interface has very limited options for the user to interact. That way, it is unlikely for the user to get confused. This approach complies with the design principle of constraints.

Due to the unorthodox ultrasound communication technology, we implemented an introduction. These introductory pages appears when the user starts the application. The top screenshots in Figure 1 show two views of this introduction. These pages, designed as a carousel, give a brief overview of the ultrasound communication technology. For usability reasons, we added appropriate signifiers as dots beneath the tab indicating the active page.

To ensure discoverabilty, all possible functionalities must be visible to the user and the current status of the application apparent. All points of interactions were realised with native Android controls (i.e. buttons) to quickly access them. Furthermore, the status of the application is indicated by a monkey icon prominently placed in the center of the main screen as illustrated in the two bottom screenshots in Figure 1. This monkey icon shows to the user clearly if the app is listening for ultrasound triggers and is ready to execute actions or not.

It is important, that the user has constant knowledge about the state of the application. Aside of that, the underlying ultrasonic communication technology should be transparent. Thus, the app provides clear feedback and shows a conclusive message whenever it detects an ultrasound trigger (see Figure 1, bottom right). This helps shaping a conceptual model of

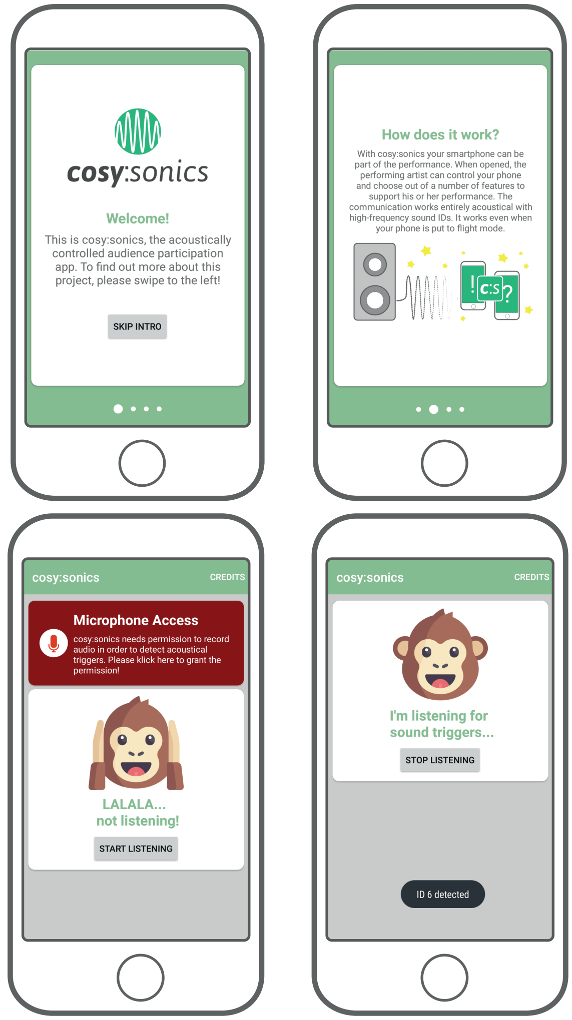


Fig. 1. App screenshots of the first two introduction views (top) and two screenshots of the ultrasound receiver views (bottom).

how a user thinks the application works which is critical to initiate technology reflection.

The *EatCoin* mining views have more content and therefore a higher density of objects. Figure 4 (bottom) shows the views of the *EatCoin* mining module. The controls in these views are grouped in blocks to resemble their affiliation and relationships (see Figure 2). The first view is the introduction to this module where the user can start the mining process (see Figure 4, bottom left). The upper half of the second screen enables interaction with the media player (see Figure 4, bottom center). The user may start, pause or rewind the track as necessary. The bottom set of controls provides options to answer the puzzle. Upon finishing the puzzle, a voucher of the mined *EatCoin* with a hash- and QR-code of the user's performance will appear (see Figure 4, bottom right). The user may save or dismiss this voucher. Saved *EatCoins* will be visible on the main screen and can be fetched at any time.

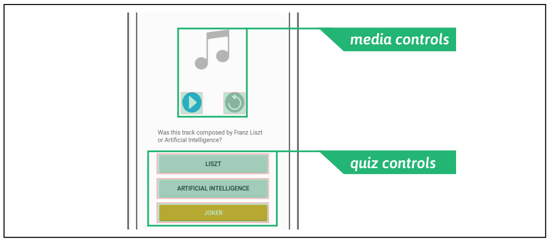


Fig. 2. Controls are grouped in blocks.

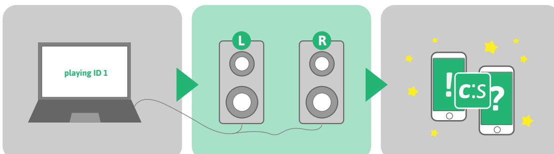


Fig. 3. A schematic overview of the communication setup.

B. Implementation

The major part we re-used from the open source app *Poème Numérique* [17] concerns the development platform itself and the ultrasound communication technology. As development platform we also used Xamarin Forms [23] which provides a framework to develop cross-platform applications for the major mobile operating systems.

Concering the ultrasound technology we adapted, the available implementation offers 17 distinguishable sound IDs to trigger certain events on nearby smartphones. One ID consists of two simultaneously played high frequency sounds in the range between 18 kHz and 20.7 kHz. Thus, a separate computer with a stereo sound system is required to transmit ultrasound triggers targeting nearby smartphones. If a mono system with one speaker is used, the two high frequency sine waves can interfere on a single speaker membrane resulting in different, not detectable sound IDs.

Figure 3 illustrates a schematic overview of the communication. Furthermore, the original implementation uses a synchronization ID to reduce false positives. A detailed documentation of this ultrasound technology is available in the work of Bartmann [17].

IV. The Final Prototype

The final app is a fully functional prototype to be used in a demo show case in various ways. The main objective is to highlight the rather unorthodox technology of communicating via high frequency sound and the vast amount of energy consumption used for Bitcoin mining. We distributed the *cosy:sonics*[1] prototype in the Google Play Store [24] to ensure that participants at the test events are able to quickly download and easily install the application. We did not publish an iOS app for the evaluation within this study. It would habe been possible due to the cross-platform development of Xamarin, but we could not test the iOS version thouroughly for these public trials.

A. Startup Screens

At startup, the *cosy:sonics* app provides a short overview of how the app works, information about the event it is used in, and a recommendation to switch the smartphone to flight mode for the best experience (see Figure 4, upper section 'startup and introduction'). The latter hint with the flight mode intends to make users aware of the fact that their phones are disconnected from all networks but still 'reachable' through ultrasound.

B. The Main Screen and Triggerable Events

The centerpiece of the main screen is the monkey icon that indicates the current status of the app (see Figure 4, middle section 'main screen and triggerable events'). If the monkey covers his ears, *cosy:sonics* is not listening for sound triggers. However, if the app is actively recording and listening for sound triggers, the monkey smiles with his ears uncovered. A text label underneath the monkey provides additional information about the status. The user can control the status of the app by pressing the Start/Stop Listening button. On startup, this button is highlighted in orange to not overlook it.

The app incorporates nine events to be triggered by receiving the respective ultrasonic frequencies. Figure 4 illustrates these nine triggerable events in the middle section on the right. In particular, these are four different screens with combinations of blinking in different colors and vibrating: (1) plain blinking, (2) blinking in the colors of the rainbow, (3) plain vibrating or (4) vibrating in the rhythm of a heartbeat. Additionally, the app provides two information slides to be triggered and shown on the users smartphone. These slides aim to raise awareness for the substantial energy consumption of Bitcoin mining. Figure 4 shows such an information slide at the bottom on the left. Finally, the *EatCoin* module opens as described next, when the respective sound ID is received.

C. The EatCoin Module

In the process of Bitcoin mining, miners have to show a Proof-of-Work by using computational power to solve specific tasks that can only be solved in a brute-force manner. This process uses substantial amounts of energy and its reasonableness is controversial. As an analogy to that, we introduce *EatCoin* mining, that consumes various amounts of the user's time instead of computational power. Three ultrasound IDs can trigger this module to start a different puzzle. Any puzzle includes a different amount of questions and therefore take varying amounts of time to solve.

For the puzzle the user listens to a number of samples of classical music pieces and has to decide, whether the piece was composed by a human composer or by artificial intelligence. The music examples are based on *Kulitta*, a framework for automated music composition [25]. At the end of the puzzle, the user receives an *EatCoin* as reward, regardless of the

[1] Note that the name of the application is a fusion between *COSY* indicating the responsible research group "Cooperative Systems" at the University of Vienna, and *sonics* to highlight the ultrasonic communication approach.

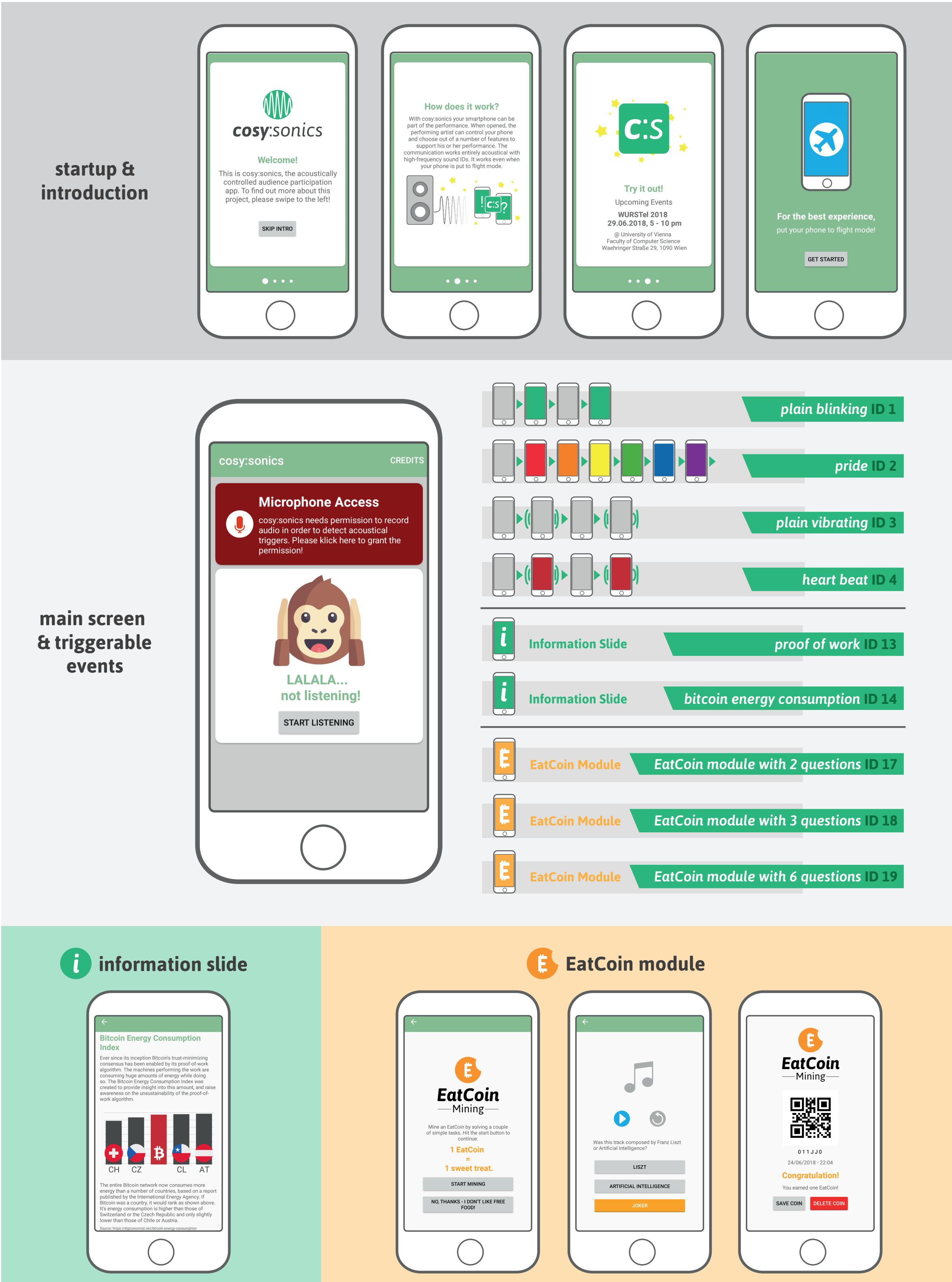


Fig. 4. cosy:sonics prototype overview

Modell-No	OS	API level
Samsung Galaxy S3 (GT-I9300)	Android 6.0.1	23
Samsung Galaxy A3 (SM-A310F)	Android 5.1	21
Samsung Galaxy S6	Android 7.0	24
OnePlus 6	Android 8.1	27

TABLE I
THE FOUR TEST DEVICES.

Fig. 5. Impressions of the demo and students using the *cosy:sonics* app

performance at the puzzle. Crucial to the process of mining *EatCoins* is to consume the users time and not whether answers are correct or incorrect.

After completing the quiz, the user will see a voucher for one *EatCoin* with a summary of his performance and a QR-Code for verification. If the user chooses to save the *EatCoin*, this voucher is saved persistently within the app. Saved *EatCoins* will appear on the main screen. The voucher can be reopend by clicking on the button.

V. EVALUATION

For evaluation, we conducted functional tests, an in-situ evaluation, and a focus group with experts. Right after the app development, we tested the prototype's functionality on different smartphone devices. For an in-situ evaluation, we demonstrated and used *cosy:sonics* during a get together event for students and faculty members at the University of Vienna. Finally, we discussed the prototype with an expert group of the Austrian Federal Ministry for Transport, Technology and Innovation.

A. Functional Tests

The functional tests included installing and testing the prototype on four different smartphones. Acoustic triggering was executed with off-the-shelf speakers (Logitech Z323) in a quiet, 20 square meter sized room. The smartphones were positioned approximately 2 meters away from the speakers. Varying acoustical features of other rooms were not considered in this test. For the actual test, all implemented sound IDs were triggered consecutively. All devices successfully detected the correct ultrasound IDs. We tested *cosy:sonics* on four different Android-based smartphones as listed in Table I. The app worked without any noticeable problems on all test devices.

B. In-situ Evaluation

We conducted an in-situ evaluation to explore the potential of the *cosy:sonics* app prototype and to collect basic data from useres. The evaluation took place during the annual *Wiener Usability Research Symposium in Telecommunications (WURSTel)* at the computer science faculty of the University of Vienna on 29$^{\text{th}}$ June 2018. We did the live demo after the official lectures and presentations in a casual lobby environment. Interested students were invited to join in an adjacent but separate room and were asked to install the prototype. Figure 5 shows the setting and students using the app.

We demonstrated all available functionalities by triggering all nine ultrasonic sound IDs. For each triggered ID, we gave users enough time to understand the content and explore the functionality. By doing so, we wanted to ensure that all participants had enought time to fully experience the functionalities and understand the content of *cosy:sonics*. The information slide components within the app were used to give information about the Proof-of-Work concept of Bitcoin and the vast energy consumption that comes with it. The *EatCoin* mining module was used as an analogy to that, using time of the user instead of energy for the mining process.

After the live demo, all participants were asked to fill out a short questionnaire. Table II lists the six questions and summarises the results of this survey. Finally, every user could trade the *EatCoin* mined with *cosy:sonics* for a piece of chocolate, i.e. a so-called *Mozarttaler*, a traditional Austrian tourist gift.

C. Expert Focus Group

In addition to the in-situ evaluation of *cosy:sonics*, we presented and discussed the prototype in an expert focus group consisting of officials from the Austrian Federal Ministry for Transport, Innovation and Technology as well as researchers from the University of Vienna. The presentation started with of a brief theoretical introduction and a short talk about the context and idea behind *cosy:sonics*. The presentation concluded with a live demonstration where the participants could experience the app with their own smartphones.

Survey		n = 15	
No	**Question**	**Yes**	**No**
1.	**I am a computer science student.**	8	7
2.	**I knew that high-frequency sound can be used as a medium of communication.**	10	5
3.	**If so, which applications do you know?** *Excerpt of answers:* *advertising, tracking, bats, therapeutical applications, medicinical diagnostics, short range communication*		
4.	**I was aware of the fact that bitcoin mining consumes a lot of energy.**	8	7
5.	**cosy:sonics worked without problems on my device.**	10	5
6.	**If it didnt, please provide a short description of the problem and the device you used.** *Some phones had troubles detecting the sound triggers. Most participants were able to sort the problem out by getting closer to the speakers or position themselves more centrally in the room, respectively.*		

TABLE II
QUESTIONS AND SUMMARISED RESULTS OF THE SURVEY.

In the aftermath, we openly and informally discussed the suitability of *cosy:sonics* to be used as social event access control mechanism at a big IT event to be held in Vienna later in 2018. Several members of the group were concerned about the fact that potential users have to download and install the app to participate in the experiment. They were sceptical, if enough visitors of the conference were willing to overcome that barrier. Another point of criticism was the fact that the app combined too many different topics concerning new technologies. The experts were afraid this could lead to confusion for the users and that it would likely be better to focus entirely on one issue.

VI. DISCUSSION

The subsequent discussion of the findings and how they relate to the literature referenced earlier is structured across the three main points of criticism that evolved.

A. Ultrasonic Communication

The communication technology did not work flawlessly for one third of the subject group. While it was not possible to perform a definite analysis of the cause of the problems, it likely has to do with the participants' positioning in the room. To detect an ultrasound trigger, two distinct ultrasounds played by two different speakers must be heard by the smartphone. Being out of the range of one of the speakers could lead to problems. According to our observations and user feedback, it helped them to receive ultrasound triggers when they moved closer to the speakers or positioned themselves more centrally in the room. Aside of that, local conditions may influence room acoustics and lead to further problems rendering the communication technology flaky. The authors of *Poème Numérique* observed similar problems when using ultrasound communication [16], [17].

Another problem concering the reliability is the lack of general standards for smartphone microphones. Different manufacturers use diverse hardware and software, making results of audio processing vary. Kardous et al. [12] conducted extensive experiments comparing the performance of several sound measurement applications on smartphones and found that results very much depend on the manufacturer of the phone aside of the used software.

Furthermore, it may be possible that smartphone manufacturers will limit the range of frequencies that can be detected by their phone's microphones in future [11] or that users utilise software to prevent undetected communication [5]. This could interfere with ultrasound communication technology in general independent from the use case and even when people are fully aware of the fact that there smartphones are triggered remotely using ultrasound as in our demo scenario.

B. Native versus Web App

A conceptual point of criticism was that *cosy:sonics* is a native application. Experts from the ministry for transport, technology and innovation feared that download and installation of the app posed too big of an obstacle for many of the visitors of their conference. Indeed, today's users tend to install fewer native apps and spend less time using them than a couple of years ago [26]. We could not ask all participants of our event. But given the fact that we just had 15 participants and there were over hundred at the event overall indicates, that the installation of the app could have been a barrier for many of them.

Another solution would have been using a web app. We could not realise *cosy:sonics* as a web-based application, however, due to the lack of access to native sensor APIs. Aside of that, we wanted to create an application that could be downloaded in advance and that works entirely without the need of internet access, even with the phone put to flight mode. While the possibilities of web-based apps have considerably improved over the last few years, apps that require direct access to certain sensor APIs still have to be implemented natively [27]. Thus, we implemented *cosy:sonics* as native app in the need to access the microphone API to detect ultrasonic sound triggers. With our insights, we contribute to an ongoing discussion within mobile development research to re-think the taxonomy of web, native and hybrid mobile development [28].

C. Critical Technology Reflection

With *cosy:sonics*, we aimed to study and raise critical technology reflection for two independent topics: (1) the possibilities of inaudible communication through ultrasound and (2) the vast energy consumption of Bitcoin mining. The experts reviewing the application noted, that this could lead to confusing or overwhelming the user. They recommended to define and focus on one key topic. However, the participants of the live demo at the in-situ study did not appear to be overwhelmed by the mix of two different topics according to the results of the survey.

It is rather striking to see the results of asking our participants about the technology awareness. One third of our study participants did not knew about the possibiliy to use ultrasound as medium for communication. Even more, nearly half of them did not know about the vast energy consumption Bitcoin mining produces. This underlines the importance to foster critical technology reflection and raise awareness for unknown possibilities of new technologies. However, we do not know on the basis of our results whether our immersive technology-centred approach enhances the effectiveness of reflection processes among students [6].

VII. Conclusion

We developed *cosy:sonics*, a mobile app to trigger technology reflection and raise awareness for both, inaudible ultrasound communication and the vast energy consumption of Bitcoin mining. For evaluation, we tested the application in-situ at a live demo with students and discussed it with experts from the domain. The results show, that there are accumulated needs to foster critical reflection among students. A considerable amount of students did not know about the possibilities of ultrasound communication and the sideeffects of cryptocurrency mining. The experts were mainly concerned about overwhelming people when confronting them with more than one challenging topic and the problems with native mobile apps preventing people from using them.

Future work has to be done to improve the ultrasonic communication technology within the app itself. It needs improvement especially regarding its reliability. Aside of that, possibilities to create a web-based application should be analysed. For the technology reflection part, we could further measure if and how the *cosy:sonics* app concept prompts the users to reflect on technology. Finally, a study with a larger audience and a sample aside from students could bring valuable insights.

Acknowledgment

This work was realised in collaboration with the VRVis Competence Center. VRVis is funded by BMVIT, BMWFW, Styria, SFG and Vienna Business Agency in the scope of COMET - Competence Centers for Excellent Technologies (854174) which is managed by FFG.

References

[1] N. Confessore, "Cambridge Analytica and Facebook: The Scandal and the Fallout So Far," [Accessed 14. September 2018]. [Online]. Available: https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

[2] N. K. Lankton and D. H. McKnight, "What does it mean to trust facebook?" *ACM SIGMIS Database*, vol. 42, no. 2, p. 32, 2011.

[3] J. Lanier, *Ten Arguments For Deleting Your Social Media Accounts Right Now*. Bodley Head, 2018.

[4] D. Malone and K. O'Dwyer, "Bitcoin Mining and its Energy Footprint," in *25th IET Irish Signals & Systems Conference 2014 and 2014 China-Ireland International Conference on Information and Communities Technologies (ISSC 2014/CIICT 2014)*, 2014, pp. 280–285.

[5] M. Zeppelzauer, A. Ringot, and F. Taurer, "SoniControl - A Mobile Ultrasonic Firewall," in *MM '18*, Seoul, Rep. of Korea, 2018, pp. 1–3.

[6] K. Kori, M. Pedaste, Ä. Leijen, and M. Meots, "Supporting reflection in technology-enhanced learning," *Educational Research Review*, vol. 11, pp. 45 – 55, 2014.

[7] P. Slovák, C. Frauenberger, and G. Fitzpatrick, "Reflective Practicum: A Framework of Sensitising Concepts to Design for Transformative Reflection," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, 2017, pp. 2696–2707.

[8] X. Lin, C. Hmelo, C. Kinzer, and T. Secules, "Designing technology to support reflection," *Educational Technology Research and Development, 1999, Vol.47(3), pp.43-62*, 1999.

[9] N. K. Lankton and D. H. McKnight, "Do People Trust Facebook as a Technology or as a "Person"? Distinguishing Technology Trust from Interpersonal Trust," in *Conference: Learning from the past & charting the future of the discipline. 14th Americas Conference on Information Systems, AMCIS 2008*, Toronto, Ontario, 2008.

[10] T. Hsu, "For Many Facebook Users, a Last Straw' That Led Them to Quit," 2018, [Accessed 14. September 2018]. [Online]. Available: https://www.nytimes.com/2018/03/21/technology/users-abandon-facebook.html

[11] L. Deshotels, "Inaudible sound as a covert channel in mobile devices," in *Proceedings of the 8th USENIX Conference on Offensive Technologies*, ser. WOOT'14, Berkeley, CA, USA, 2014, pp. 16–16.

[12] C. A. Kardous and P. B. Shaw, "Evaluation of smartphone sound measurement applications," *The Journal of the Acoustical Society of America*, vol. 135, no. 4, pp. EL186–EL192, 2014.

[13] D. Arp, E. Quiring, C. Wressnegger, and K. Rieck, "Privacy Threats through Ultrasonic Side Channels on Mobile Devices," in *Proceedings - 2nd IEEE European Symposium on Security and Privacy*, 2017, pp. 35–47.

[14] V. Mavroudis, S. Hao, Y. Fratantonio, F. Maggi, C. Kruegel, and G. Vigna, "On the Privacy and Security of the Ultrasound Ecosystem," in *Proceedings on Privacy Enhancing Technologies*, no. 2, 2017, pp. 95–112.

[15] M. Hirabayashi and K. Eshima, "Sense of Space : The Audience Participation Music Performance with High-Frequency Sound ID," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2015, pp. 58–60.

[16] F. Kayali, C. Bartmann, O. Hödl, R. Mateus-Berr, and M. Pichlmair, "Poème Numérique: Technology-Mediated Audience Participation (TMAP) Using Smartphones and High-Frequency Sound IDs," in *INTETAIN 2016*, 2017, vol. LNICST 178, pp. 254–258.

[17] C. Bartmann, "Exploring audience participation in live music with a mobile application," Master Thesis, Vienna University of Technology, 2016.

[18] B. Smagowska and M. Pawlaczyk-uszczyska, "Effects of ultrasonic noise on the human bodya bibliographic review," *International Journal of Occupational Safety and Ergonomics*, vol. 19, no. 2, pp. 195–202, 2013.

[19] Digiconomist.net, "Bitcoin Energy Consumption Index," 2018, [Accessed 14. September 2018]. [Online]. Available: https://digiconomist.net/bitcoin-energy-consumption

[20] L. Kugler, "Why cryptocurrencies use so much energy," *Communications of the ACM*, vol. 61, no. 7, pp. 15–17, 2018.

[21] J. Truby, "Decarbonizing Bitcoin: Law and policy choices for reducing the energy consumption of Blockchain technologies and digital currencies," *Energy Research & Social Science*, no. June, pp. 0–1, 2018.

[22] D. A. Norman, *The Design of Everyday Things*, 1st ed. New York: Basic Books, 2002.

[23] Microsoft, "Xamarin forms documentation," https://docs.microsoft.com/en-us/xamarin/xamarin-forms/, [Accessed 14. September 2018].

[24] P. Oberhauser, "cosy:sonics in the google play store," https://play.google.com/store/apps/details?id=com.companyname.CosySonics, [Accessed 14. September 2018].

[25] D. Quick, "Kulitta," [Accessed 14. September 2018]. [Online]. Available: http://donyaquick.com/kulitta/

[26] T. Ater, *Building Progressive Web Apps: Bringing the Power of Native to the Browser*. O'Reilly, 2017, pp. 3–6.

[27] A. Puder, N. Tillmann, and M. Moskal, "Exposing native device apis to web apps," in *Proceedings of the 1st International Conference on Mobile Software Engineering and Systems*, ser. MOBILESoft 2014. New York, NY, USA: ACM, 2014, pp. 18–26.

[28] R. Nunkesser, "Beyond Web / Native / Hybrid : A New Taxonomy for Mobile App Development," pp. 214–218, 2018.

Promenade

A Minimalistic Sculptural Interface for Immersive Soundscape Presentation

Martin Rumori
Institute of Electronic Music and Acoustics
University of Music and Performing Arts
Graz, Austria
Email: rumori@iem.at

***Abstract*—*Promenade* proposes an interface for the semi-personal, interactive presentation of soundscapes that allows for the seamless integration of visual material. It combines the excitation of a flat surface using transducers with transaural cross-talk cancelled stereophonic projection. The unfolding of a wide spatial image within the sweet spot is contrasted by a collapsed, localised impression outside. This inherent interactive behaviour may be further extended by means of a distance sensor.**

***Promenade* emerged from an experimental, arts-based research project. It therefore targets contexts that feature custom audio content rather than aiming at general-purpose applications requiring optimal acoustic properties.**

I. Introduction

Exhibitions or other presentations of distinct objects often call for an unobtrusive sound projection means that affords the visitors an immersive, mostly individual listening experience. Achieving such an experience requires spatial capabilities of the chosen projection technique.

In many such cases, headphones are used. Spatially rich sound material may be projected in particular with binaural signals. A major side effect is the inherent reconfiguration of our multi-modal perception, as the surrounding auditory domain is largely replaced by the projected one. Additionally, headphones may be considered as cumbersome or unhygienic in public exhibitions. Multichannel loudspeaker setups, in contrast, do not require physical engagement with the projection medium. However, they do not facilitate individual listening situations and they have to be mounted in an often complex constellation around the target listening area.

Promenade is a minimal auditory display prototype that relates to such issues. It is important to notice that *Promenade* has not been developed to overcome the limitations of other projection means. Instead, it emerged from a series of experiments in an arts-based research context as part of a sound installation with a focus on aesthetic listening experience. Due to its simplicity, its purpose and the idiosyncratic technology it incorporates, *Promenade* is not a universal, general-purpose spatial sound projection system but it manifests certain limitations. However, it might provide an appropriate and elegant solution in some of the circumstances described above. *Promenade's* qualities with respect to those circumstances are regarded retrospectively here in order to explore further possible fields of application and obvious modifications for different forms of appearance.

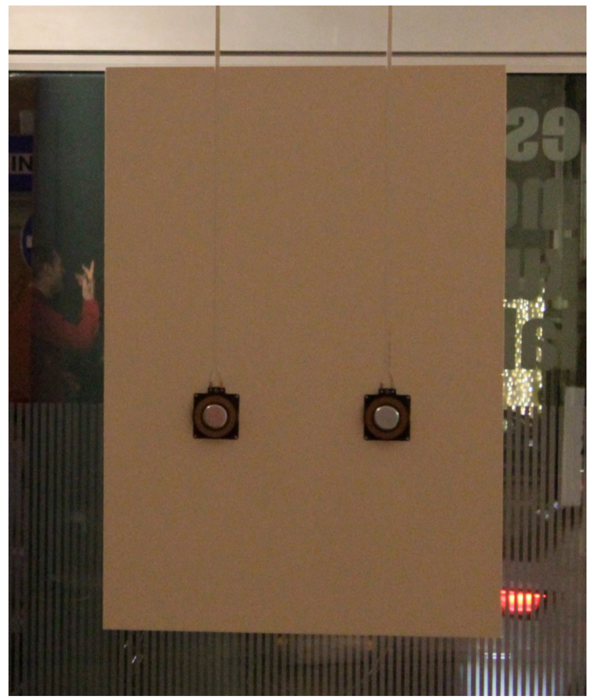


Figure 1. *Promenade* sound projection device.

This paper describes the basic idea of *Promenade* and its originating context (section II), followed by introducing the two main techniques involved, transaural crosstalk-cancelled sound projection (section III) and the principle of flat panel loudspeakers (section IV). Finally, some findings that arose from practical experience are discussed, along with possible modifications and extensions (section V).

II. Promenade

The *Promenade* sound projection device consists of a flat surface of variable size with two electrodynamic exciters mounted on the back (Figure 1). Audio signals are fed to the exciters such that the surface acts as a loudspeaker membrane. The physical properties of the surface material lead to a certain radiation characteristics. The directional effect is further emphasised by crosstalk-cancelled stereophony, which creates a large auditory canvas that is much wider than the projection panel (see section III). The effect occurs at a central listening position, facing the device from a distance of a few ten centimetres up to about one and a half metres. As soon as the sweet area is left, the spatial auditory image collapses and sound seems to be restricted to the panel surface. Although both exciters are mounted to the same surface, their signal differences remain intact to a sufficient extent for providing the effect of crosstalk-cancelled projection.

A. Physical construction

Promenade has been prototyped with two materials and in two sizes so far. The first series uses finnboard, a lightweight type of cardboard, with the dimensions 70 by 100 centimetres (portrait) and 3 millimetres thickness. The second, more compact series is made of screenboard, a more dense type of cardboard with white cover layers, 43,5 by 64 centimetres and 2 millimetres thick. Other promising materials that have been explored include ply wood of a few millimetres thickness, and KAPA, a compound material combining a polyurethane foam core with cardboard cover layers.

The smaller prototype uses Visaton EX-60 R exciters (60 millimetres in diameter), while the larger is equipped with the more powerful Visaton EX-80 S (square shape, 80 millimetres edge length). White flat cable serves for feeding the exciters and for suspending the device. Both the exciters and the cables are glued to the panels. The projection devices are fed by small standard amplifier boards installed separately.

In the installation, the projected signals were generated from a BeagleBone Black computer board equipped with the Bela audio cape.[1] Sound programming and convolution for crosstalk cancellation was performed in SuperCollider.[2] As cancellation filters, the so-called BACCH filters by Edgar Choueiri [1] were taken from the jconvolver convolution engine bundle by Fons Adriaensen.[3]

B. Sound installation and artistic context

The sound projection device has been conceived as part of a sound installation within the arts-based research project *Promenade. Anekdote, Alltag, Abbild (Promenade. Anecdote, everyday life, image),* hence its name.[4]

Seven playback canvases (five in a later version, Figure 2) were distributed in the exhibition space, each suspended from the ceiling. The projection surface was left blank, without any further visual additions. Each of the devices played back a soundscape composition based on field recordings. Listeners were invited to explore and 'visit' the different soundscapes by moving in space, approaching the projection devices one by one and 'entering' the recorded space by positioning themselves in the sweet area where the crosstalk-cancelled projection effect unfolds.

The installation seeks to combine a global sonic environment that reflects the simultaneity of seemingly unrelated

Figure 2. *Promenade* installation, SMC 2018, Limassol, Cyprus.

structures, processes, and realities around us, with the singularity of particular spots in space that, when located, allow for an exclusive access to individual layers of this environment. The crosstalk-cancelled projection technique restricts these spots to clear-cut boundaries but on the other hand it potentially 'elevates' each layer's spatiality as a sublime, temporary overlay to the global environment.

The installation *Promenade* serves as an experiential starting point for further investigations into the fields of anecdotal music, narrative identity construction and strollology. All of these areas may be reflected in conjunction with aesthetic experience in our everyday life auditory surrounding and its technical mediation. Departing from a workshop within the project *Promenade,* a publication is currently prepared that summarises the described approach and the ongoing process.

III. Transaural crosstalk-cancelled sound projection

A. Working principle

Transaural crosstalk-cancelled sound projection, often abbreviated XTC, has been developed for playing back binaural signals, which normally require headphones, on a pair of loudspeakers (see, e. g., [2, pp. 283–326]). While headphones isolate the sound propagation paths from each of the earphones to the corresponding ear, respectively, loudspeaker projection involves transaural crosstalk: the signal radiated from the left ear also reaches the right one and vice versa. However, the playback of binaural signals requires the separate presentation of each channel to the corresponding ear.

Transaural crosstalk cancellation is based on time-shifted anti-phase signals emitted from the respective opposite loudspeaker that shall arrive at the near-side ear at the same time as the unwanted crosstalk signal from the far side speaker in order to cancel it out. Obviously, the cancellation signal causes a crosstalk at the far side ear, too, which in turn is eliminated by another correction signal emitted from the opposite speaker. Several iterations of this process lead to the design of a recursive filter network that ultimately seeks to ensure an isolated projection of the ear signal pair [3].

[1] http://bela.io
[2] http://supercollider.github.io
[3] http://kokkinizita.linuxaudio.org/linuxaudio/
[4] https://esc.mur.at/de/projekt/promenade-anekdote-alltag-abbild

B. Binaural vs. stereophonic signals with XTC

As mentioned above, transaural crosstalk cancellation has been developed for the projection of binaural signals on loudspeakers, that is, ear signals involving interaural time (ITD) and level differences (ILD) as well as elevation-dependent spectral cues (see [4]). Nevertheless, XTC filters are also used for the playback of ordinary loudspeaker stereophonic signals. The latter application is known as *Ambiophonics* in consumer entertainment contexts [5].

Experimental findings of the *Promenade* project with respect to stereophonic signals are discussed below (Section V).

IV. Flat panel loudspeaker projection

The excitation of flat, solid matter surfaces as an alternative loudspeaker working principle has been investigated during the past decades (see, e.g., [6]). Major advantages are the appealing form factor requiring dramatically reduced mounting depth, and the potential to create diffuse sound fields due to the low directivity even at high frequencies.

One important factor for exciting flat panels is the mounting location of the exciter. Depending on the application, the physical properties and the size of the panel, a basic question is whether to exploit or to avoid the direct excitation of predominant modes. Often, an out-of-centre location avoiding low integer partition of the surface dimensions is recommended for a more balanced response. In *Promenade,* exciters were mounted according to the golden section as an experimentally found trade-off of acoustic and visual qualities.

Apart from the construction of high-end loudspeakers, electrodynamic excitation of manifold objects for sound projection is frequently applied in custom presentation media and in sound art (see, e.g., *SHHH SHHH SHHH*[5] by Amanda Dunsmore, 2008, or *Touched Echo*[6] by Markus Kison, 2007). As described above, the exploration for *Promenade* departs from the latter. In such realms the presented sound material is usually created or adapted especially for the specific playback medium. Consequently, the projection device does not need to meet the same criteria as general-purpose loudspeakers, especially with respect to linearity and neutrality, because such deviations can be corrected in the adaption process. This largely facilitates achieving a beneficial result with relatively low-cost equipment.

A special technical issue of the *Promenade* device is that only one panel is used for projecting two audio channels at the same time. The crucial question is whether the crosstalk cancellation cues are propagated to a sufficient extent for achieving a similar effect as with two small-spaced loudspeakers or whether the inter-channel crosstalk in the panel degrades the cues too much. A similar technique has been investigated for the application of Wave Field Synthesis (WFS) using even more than two exciters on the same panel, coined Multi-Actuator Panels [7]. Later, the method has been generalised for creating multiple sound sources with a single panel [8].

[5]http://www.lit.ie/Dunsmore/linz/kunstraumwindows.html
[6]http://www.markuskison.de/touched_echo.html

The experimental artistic context of *Promenade* did not pursue a scientific approach to such a technique for transaural sound projection. In multiple iterations, the used combination of materials, dimensions and mount points of the exciters have been adjusted such that the desired effect could be achieved sufficiently well. A validation by acoustic measurements, however, has not yet been carried out.

V. Discussion

The presented sound projection device shows an appropriate performance within its frame of origin, the *Promenade* arts-based research project. There, it allows for experiencing the often surprising, sometimes uncanny effect of crosstalk-cancelled sound projection and supports the exploration and reflection of aesthetic experience in multi-layered spatial auditory environments.

Beyond the attached context, the principles exploited in *Promenade* may be applied to other use cases. In the following, some findings and possible directions are discussed, such as the use of stereophonic signals, the integration with visual presentations, extensions towards interactive applications, and other form factors, materials, or objects.

A. Crosstalk cancellation of stereophonic signals

When developing the *Promenade* sound installation with field recording material and XTC projection, it was found that coincident stereophonic signals involving only level differences but none of phase provide a more consistent and stable spatial image than recordings in spaced omnidirectional stereophony that is based on phase differences. There is no well-researched explanation for this observation. It is likely that both the listening conditions and the projection device are not optimal for propagating phase differences to the listener's ears. In particular the fact that both exciters are working on the same projection surface may degrade the fidelity of phase reproduction, although a similar observation was made using two separate loudspeakers.

Overall, the discovery that intensity stereophony is well suited to span the extremes of the projection base – from ear to ear using headphones, nevertheless crossing the listener's head, from speaker to speaker in an ordinary stereophonic setup, and along a well externalised, virtual stereo base of about 150 degrees opening angle using XTC projection – is most valuable for the production of appropriate material. Following from that, field recordings were carried out in mid/side stereophony, which allows for further benefits with respect to adaptation and interaction (see subsection V-C).

B. Visual integration

In the *Promenade* sound installation, major emphasis was put on listening to mediated auditory environments. Therefore, the metaphor of a canvas was picked up that could have shown a landscape but instead projected a soundscape while remaining visually 'silent'. However, the surface of the projection medium may be easily used for a seamless integration of visual and auditory material. In the simplest case, the cardboard material may be printed or laminated. The backside may be used as well, restricted by the mounted exciters.

Alternatively, the panel may serve as a video projection screen, including three-dimensional technology involving polarised or shutter glasses. The projection angle and distance have to be envisaged such that the spectator is not disturbing the projection but may still position herself appropriately with respect to crosstalk-cancelled sound projection.

Using stereophonic signals, auditory events can be evoked for the listener anywhere in the horizontal dimension within the borders of the canvas and beyond. Therefore, horizontal audio-visual congruence is well achievable, for example, for annotating visual items in maps or alike. Psychoacoustic effects may be exploited for suggesting vertical correspondences. Depending on the stiffness of the projection panel, the location of the exciters on the panel is audible in certain frequency bands. In this case, a second pair of exciters may serve as a panning means in the vertical dimension, at least to a certain extent.

Generally, any visual presentation that favours a fixed, frontal and relatively close viewpoint can be expected to work well in combination with a *Promenade* like sound projection device.

C. Interactivity, tracking, sensors

From the beginning, the *Promenade* device has been conceived with an additional optical distance sensor such that the playback can react interactively, depending on whether, how close and for how long a listener is present in front of the canvas. For the installation versions involving several projection panels, actual use of the sensors has been dismissed so far. It turned out to be more desirable in the first place to explore the field of sweet spots distributed in space which, when being entered, change the listening experience dramatically. This effect may be regarded as an inherent 'interactive' behaviour of the projection dispositif, although no technical reactivity is involved.

Nevertheless, involving the sensor bears several interesting use cases. If multiple projection panels are present, the sensors may serve as a simple tracking device in order to activate or emphasis one or more panels rather than others, both in a decentralised (each panel controls itself without influencing others) or a centralised manner. Furthermore, the spatial arrangement of sound material could be extended towards an enterable, interactive radio play. Different individual listeners cannot be distinguished, though, which would require more sophisticated tracking means.

The sensor information may also affect the projection technique itself. In the *Promenade* installation, gradual transitions take place from monaural signals (centered in the projection panel) over ordinary stereophony (located roughly in between the exciter positions, thus a partition of the panel) to crosstalk-cancelled projection. As mid/side-stereophony is used, modulating the contribution of the side signal allows for further controlling the perceived width of the resulting sonic image. Such transitions may be triggered or emphasised by evaluating the listener's distance or overall presence via the sensor.

D. Other form factors, materials, or objects

Theoretically, combining two-channel electrodynamic excitation of solid matter with transaural crosstalk cancellation is possible with various objects and materials of any form or extension. For example, architectural prototypes, engineering models or basic commodity might be enhanced by an integrated auditory layer that invites for spatial immersion.

The determining factors remain those that have been consulted while conceiving *Promenade:*

- the physical properties of the excited material with respect to sound propagation,
- appropriate mounting locations for the exciters and fitting form factors of the object and the exciters, and
- sufficiently isolated radiation of both channels for achieving the desired effect of crosstalk cancellation.

VI. Conclusion

In this paper, the *Promenade* spatial sound projection device has been presented, along with its originating context, a sound installation as part of the arts-based research project of the same name. The device has been presented as combining the two techniques of transaural crosstalk cancellation for inviting a rich spatial auditory experience with the excitation of solid objects for an unobtrusive appearance, differing from that of ordinary loudspeakers. The presentation form of a canvas, though not visually occupied in the original installation, invites for manifold applications that involve audio-visual correspondence. Further extensions include the use of distance sensors for interactive functionality, both with respect to the presented narrative and to the evoked spatial impression.

Acknowledgements

The research project *Promenade* is supported by the Vice Rector's office of research at University of Music and Performing Arts Graz, the Cultural Office of the City of Graz, and ESC Media Arts Laboratory, Graz.

References

[1] E. Y. Choueri, "Optimal crosstalk cancellation for binaural audio with two loudspeakers," 2010. [Online]. Available: https://www.princeton.edu/3D3A/Publications/BACCHPaperV4d.pdf

[2] B. Xie, *Head-Related Transfer Function and Virtual Auditory Display*. Plantation: J. Ross, 2013.

[3] R. Glasgal, "360° localization via 4.x RACE processing," in *123rd AES Convention*. AES, 2007.

[4] J. Blauert, *Spatial Hearing*. Cambridge Mass./London: MIT Press, 1997.

[5] R. Glasgal, "Ambiophonics: The synthesis of concert-hall sound fields in the home," in *99th AES Convention*. AES, 1995.

[6] G. Bank and N. Harris, "The distributed mode loudspeaker – theory and practice," in *13th Microphone & Loudspeakers AES Conference*. AES, 1998.

[7] M. M. Boone, "Multi-actuator panels (MAPs) as loudspeaker arrays for wave field synthesis," *Journal of the Audio Engineering Society*, vol. 52, no. 7/8, pp. 712–723, 2004.

[8] M. C. Heilemann, D. Anderson, and M. F. Bocko, "Sound-source localization on flat-panel loudspeakers," *Journal of the Audio Engineering Society*, vol. 65, no. 3, pp. 168–177, 2017.

Session 2: Augmented Reality & Visualization

Potential of Augmented Reality in the Library

Bernadette Baumgartner-Kiradi, Michaela Haberler, Michael Zeiller
Information Technology and Information Management
University of Applied Sciences Burgenland
Eisenstadt, Austria
1984bbk@gmail.com | michaela.haberler@gmx.at | michael.zeiller@fh-burgenland.at

***Abstract*—Augmented reality (AR) applications are getting popular since they integrate the real world and the virtual world. This paper focusses on AR applications for public and scientific libraries. There exist some projects and sample implementations of AR apps specially designed for libraries, but they are seldom found in practice. To identify whether there is potential for AR apps to be applied in libraries, the results of a qualitative study that has been performed among librarians working in public and scientific libraries in Austria and experts in augmented reality are presented. Searching for media, navigating to the correct location and displaying ancillary information, like ratings, reviews, secondary media, links, etc., has the highest potential for users according to the experts. AR apps for maintaining the bookshelfs of a library provide real benefit for librarians and are awarded high potential as well. Guided tours through libraries using AR-based apps lightens the load of the librarians to introduce new users to the library. The feasibility of an AR app for a library is demonstrated by introducing a prototype that supports library users with additional information on media and the library itself.**

Keywords—library; public library; scientific library; teaching library; augmented reality

I. Introduction

Augmented reality (AR) bridges the gap between the real world and the virtual world – spatially and cognitively [1]. It allows to integrate the real world and the virtual world and presents information on mobile devices that directly corresponds to the physical environment. By selecting, filtering, and visualizing virtual objects context-based information can be displayed together with real world objects.

There exist several examples for various fields of applications, i.e., industry and construction, maintenance and training, the medical domain, personal information, or navigation [1] [2]. Although a number of successful applications can be found, AR is categorized by Gartner in its Hype Cycle for Emerging Technologies 2018 to be part of the *Trough of Disillusionment* [3]. Thus, it will take about another five years to reach the mass market.

Libraries are another field of applications that gets slowly into focus of AR. A few prototype applications exist that demonstrate the feasibility of augmented reality to support users and the staff of libraries. For example, the University of Applied Sciences Potsdam developed a concept and prototype for a complex AR-based app *myLibrARy* providing additional information and reviews to media as well as general information about the library [4]. LibrARi is an image-based AR app for mobile devices and AR glasses that supports users on finding their way to the desired book in the bookshelf [5]. The Miami University in Oxford, Ohio developed an AR-based app called ShelvAR that supports librarians to identify books in the wrong place and for inventory [6]. The Bavarian State Library Munich provides a location-based AR app that offers additional information to special locations, buildings, and monuments related to king Ludwig II [7].

Although there exist a number of augmented reality applications specially designed for libraries (i.e., especially for the users of libraries) their number is limited. None of them got ready for the market and some of them, like ShelvAR, even discontinued. This might be due to an immature technology, financial issues, a lack of acceptance among users or a lack of acceptance among librarians. To identify the potential of augmented reality apps in libraries we will focus on the point of view of libraries. Therefore, in this paper we will investigate whether there is a need for AR apps in public and scientific libraries, limiting the research view on Austrian libraries (section IV). To demonstrate the feasibility of an AR app for a public or scientific library a prototype supporting library users with additional information will be presented (section V).

II. Augmented Reality

A. Definition

Augmented Reality integrates 3D virtual objects in a 3D real environment in real time [8]. In AR virtual objects are superimposed upon or composited with the real world, but the user can still see the real world. Therefore, AR supplements reality, rather than completely replacing it like in *Virtual Reality* (VR) [8]. Instead of creating a completely synthetic world (i.e., virtual world) in which the observer is completely immersed, the *Mixed Reality* merges real and virtual worlds. According to Milgram and Kishino there is a reality-virtuality continuum that is related to a mixture of classes of objects. With the real environment (consisting solely of real objects) and the virtual environment (consisting solely of virtual objects) as the two opposite extrema, there is the *Mixed Reality* in between in which real world and virtual world objects are presented together within a single display [9] [10]. *Augmented Reality* is the part of the *Mixed Reality* more close to the real environment where real objects are more dominant than virtual objects.

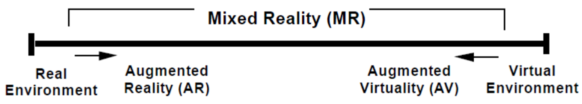


Fig. 1. Reality-Virtuality (RV) Continuum [10]

The most widely accepted definition that does not limit AR to specific technologies was proposed by Azuma in 1997 and defines Augmented Reality as “systems that have the following three characteristics:

1) Combines real and virtual
2) Interactive in real time
3) Registered in 3-D” [8]

B. AR Systems

A complete AR system requires at least three components:

- a tracking component,
- a registration component,
- a visualization component [1].

A fourth component is required to store information about the real world and the virtual world in a spatial model. The tracking component determines the location of the user in the real world. The real world model serves as a reference for the tracking component. The registration component is responsible for the alignment of coordinate systems between virtual and real objects since the virtual information has to be accurately registered with physical scene objects [1].

When using an AR system there is a feedback loop between the human user and the AR system (Fig. 2) [1]. While the user observes the AR display and controls the viewport, the system tracks this viewport and registers the pose in the real world with the virtual content. Then the situated visualizations are presented on the display [1].

Fig. 2. Feedback loop between user and AR system [1]

To be able to position virtual objects registered to real objects in 3D the relative pose – i.e., the position and orientation of the AR display relative to the real world – has to be known. Pose measurements have to be updated continuously since AR operates in real time. In AR 3D tracking refers to this process of constantly identifying the three-dimensional position or six-dimensional pose of real objects [1]. Tracking can be done using sensors exploiting physical phenomena like electromagnetic radiation (e.g., light, radio signals, magnetic flux), sound, gravity, or inertia [1] [11]. Mobile tracking uses magnetometer, gyroscope, accelerometer, or GPS. Optical tracking is based on cameras (visible light, infrared light). Optical tracking often uses markers: known patterns that are placed on the surface of real objects [1]. They are designed to make it easy to be detected in the image: square or circular shape, black border, high contrast [2]. In situations where no artificial markers can be applied, natural features of objects can be used instead, e.g., by identifying interest points, tracking edge features, or high differences in color or contrast [1] [11].

AR displays must be capable of combining the real environment and the virtual environment. An optical see-through display uses an optical element (partially transmissive and partially reflective) to combine the view of a user of the real world with computer generated images of virtual objects [1]. On a video-see-through display the real world is captured using a video camera and the image is modified electronically (using a Digital Combiner) to add the virtual objects. The combined image is then displayed on a screen [2] [1]. Based on those two fundamental principles different categories of displays can be found: head-mounted displays (e.g., smart eyeglasses), handheld displays (e.g., smartphones, tablet computer), and projective displays (e.g., head-up displays using the windscreen of cars or aeroplanes).

Additionally, users have to be able to interact with the virtual world: by movement (navigation), by selecting objects (e.g., physically grabbing a marker or pointing on a virtual object), manipulation of a virtual object (changing a parameter, e.g., turning, scaling, moving the object), input of symbols (e.g., using gesture or a virtual keyboard, or speech-driven) [2] [11].

III. Augmented Reality Applications in Libraries

There already exist a few AR applications dedicated to the use in libraries. However, most of them are prototypes ore specially designed for a certain library. This section provides an overview of the most striking AR projects for libraries. Different types of AR applications can be distinguished:

A) apps providing additional information on media for library users (including locating media in the library)
B) apps supporting librarians (e.g., for identifying books)
C) apps providing additional information on cultural assets associated with the library/archive
D) augmented books.

A. AR Apps Providing Additional Information on Media

In 2014 the University of Applied Sciences Potsdam started *myLibrARy*, a project to evaluate and explore the fields and scope of application of AR in public libraries [4] [6]. The main goal was the development of a user-oriented app for libraries with significant features related to augmented reality. Based on a user survey the app should provide features like

- managing media: search, reservation, extract, download
- navigation: finding the way to the media, virtual tour
- information on the library: opening times, contact, staff
- user service: user account, wish list, reminder, renewal
- interface to other services like bibliographic management, Wikipedia, book trade, exhibitions, events, etc.
- social media
- services outside the library like a literature walk

Some of these features are related to AR, some are classic features of library apps that could already be found elsewhere. The first prototype was developed by Metaio and was a channel in the Junaio app of Metaio [4] [12]. Media were identified by optical tracking and image recognition of the book cover. The second prototype was implemented as an independent app where media were identified by scanning the ISBN code. The main idea of *myLibrARy* was the development of Smart Libraries where smart technologies are an integrated part of the use experience. AR might be a key element that enables new knowledge due to semantic and visual contextualization of information [4] [12].

LibrARi is an image-based AR app for mobile devices and AR glasses that supports users on finding their way to the desired book in the bookshelf [5]. Since the AR app shows the direct way to the book on the display, the library can be explored interactively and users do not have to bother with classification systems any more. LibrARi offers searching, locating and navigating in physical space using a digital interface on a mobile device [5].

The University of Illinois Library developed a mobile recommender app with augmented reality features called Topic Space. By embedding optical character recognition software, the augmented reality app can recognize the signature on a book in the library and suggest relevant items that are shelved nearby. Additionally the app shows users media that are normally shelved at the location, but that are currently checked out [13].

B. AR Apps Supporting Librarians

The Miami University in Oxford, Ohio developed an AR-based app called ShelvAR intended to support librarians to identify books located in the wrong place in the bookshelf [6] [14]. Using the prototype librarians could aim the mobile device (e.g., smartphone or tablet) at the shelf and books on the wrong place will be marked accordingly (✓ or ✗). Unfortunately, due to a patent dispute the project discontinued [15].

C. Other AR Applications

The Bavarian State Library Munich developed the AR app "Ludwig II" that presents location-based services of cultural assets (e.g., special locations, buildings, monuments) related to king Ludwig II [7]. The multimedia content is either displayed on top of a map view or integrated in the live camera video.

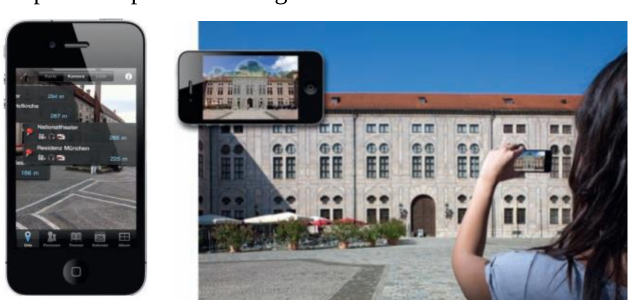


Fig. 3. AR app "Ludwig II" [7]

By tracking the pose of the mobile device location-based and context-based information (e.g., historic images, maps, letters, audio clips of contemporary witnesses) can be identified and inserted in the live video. Figure 3 left shows the camera view of the app with information on nearby points of interest. Figure 3 right shows the real time simulation of the winter garden that had been torn down after the death of Ludwig II.

D. Augmenting books

The SCARLET project (Special Collections using Augmented Reality to Enhance Learning and Teaching) developed a marker-based app (using QR codes and book covers) to support students that have to consult rare books, manuscripts and archives within the controlled conditions of library reading rooms. The AR-based app enables students to experience the best of the real and the virtual world: students can enjoy the sensory delights of seeing and handling original materials, while enhancing the learning experience by enhancing the object with digital images, online learning resources and information on the items before them and on related objects held in the library and elsewhere [16].

Like in SCARLET markers can be embedded in various kinds of books (children's books, educational books), magazines or catalogues thus allowing the reader to access additional information (e.g., various multimedia content).

IV. Evaluating the Potential of AR for Public and Scientific Libraries

There are a few AR applications available for libraries and in some of the projects presented in section III a requirements analysis has been done prior to implementing the AR app. However, none of them got ready for the market and is available for other libraries. Therefore, we will investigate whether there is potential for an AR app in libraries. We will focus both on public and scientific libraries and derive whether there is a demand for augmented reality applications [17].

Research will concentrate on opportunities and challenges when libraries use AR apps to support their users. We would like to know which types of applications and which features are relevant for users in public and scientific libraries. Special focus will be on some of the core tasks of libraries – dissemination of information and conveying media competence – to identify whether AR has a positive influence on information literacy of library users [17].

There is a range of stakeholders related to a public or scientific library. In this research we concentrate on *users* of a library, i.e., regular readers and casual readers, who are searching for media (books, journals, etc.) – either print or electronic media – and want to lend these media.

A. Method

To investigate the potential of AR apps a qualitative research method has been chosen [18]. Structured guideline interviews have been carried out with twelve experts in three different domains (four experts in each domain): augmented reality, public libraries, scientific libraries. The four experts on augmented reality are professors at universities and COOs in

AR companies in Germany, Austria and Switzerland (DACH region) and authors of books on AR. The experts on public libraries are librarians and heads of public libraries in Germany and Austria. The experts on scientific libraries are librarians and heads of libraries at universities and universities of applied sciences in Austria.

The interviews were conducted in March and April 2018 and took about 30 to 50 minutes. A mobile audio recorder was used to record the interviews. Afterwards the audio was transcribed to provide written material [19].

A qualitative content analysis according to Mayring was then applied to identify corresponding statements in the interviews [20]. To be able to analyze the material, categories had to be defined, i.e., main categories and sub-categories [18] [20]. For example, the main categories were: current offerings in libraries, technology and digital offerings, reasons for utilization of AR, reasons for not using AR, information literacy with respect to AR, AR general, AR applications, AR benefits, AR challenges, AR considerations of users, and IT affinity [17]. The text was coded according to those categories using the software MAXQDA.

B. Results

Selected and aggregated results of the conducted interviews will be presented in this section [17].

1) AR General

The AR experts assume that AR will be applied everywhere in the future because it provides support in multiple ways. Users have to get acquainted to AR apps and realize their potential, but then the transformation to its utilization will be gradually and AR apps will be used everywhere. Data goggles will be the natural device for using AR.

Currently AR is mainly applied in navigation (head-up displays), sports broadcasting (lines faded in), industry (overlay of technical documentation), logistics, construction, medicine (insert extra data during surgery), military, and gamification. Although AR is still quite new to most companies, the AR experts think that in future AR will be heavily used to visualize a lot of different issues, e.g., in architecture, marketing, content combination, and in traineeship and education.

2) Advantages of AR for Libraries

Since the discussed applications of AR in libraries are fixed to the place of the library, users have to come to the library to be able to use the AR app. Although this seems to be a trivial fact, users will be forced to visit the library – and then they will get a lot of additional information. Since AR helps to better find the way in the library, media (i.e., books) can be found much easier and faster.

Ancillary information can be conveyed much simpler, more extensive and more context specific with AR. Using 3D, information can be perceived and imagined more easily. AR can be used whenever complex information in the library shall be presented to the audience. Added value arises when the physical holdings of the library can be combined with the virtual holdings. The librarians (of both types of libraries) would like to use such combinations much more often. AR can be used to link different types of the library holdings and to show all media since not everything can be presented in the catalogue of the library. By filtering and selecting content elements more precise results can be presented to users.

Five out of twelve interviewees (1 AR expert, 3 librarians of public libraries, 1 librarian of a scientific library) mentioned that AR will help to project a modern, attractive image of the library. Using a modern and innovative technology like AR the quality of the experience in the library will be increased. With the help of AR the libraries can offer better and demand actuated service.

3) Motivation, Attention, and Information Processing

All interviewees agree that AR apps can motivate users to visit a library. However, they added that the level of motivation depends on the target group (e.g., much better with young people or technogeeks). There must be a clear benefit, e.g., saving of labor, to motivate someone to use the AR app. Motivation can be increased by improving the quality of the personal experience, e.g., fascination of technology, a game, new opportunities, or the chance to discover and experience the library in a playful way.

The interviewees also agree that using AR heightens the awareness and in this way the offerings of the library are perceived more effectively. However, the opportunity to improve the awareness depends on the target group. Users can be motivated when there is an added value. The library may track all activities and related behavior and analyze the behavior of its users anonymously.

The AR experts and the librarians of scientific libraries think that with the help of AR information can be conveyed and processed much easier. They argue that more senses are appealed. However, the librarians of public libraries disagree. Two librarians of scientific libraries point out that the amount of improvement might depend on the type of service. The other two librarians of the scientific libraries hold the opinion that we have to be aware that the provided information might not be processed properly any more because of an overstimulation.

4) Preconditions of Augmented Reality in Libraries

All interviewees agreed that using AR has to make sense and generate extra value for users. Since there is still a lack of good applications, (library) users are not aware of the added value of AR.

There are external and internal factors that influence AR in libraries. External factors might be the interior design, lighting conditions, or the internet connection within a building. Internal factors like staff or overstimulation influence the utilization of AR as well. If the app shall be used indoors, good maps with high-resolution are needed. A WIFI/WLAN network connection would be practical as well. The rooms of the library should not be too bright or too dark and the walls should not be shiny.

There are many libraries that are run by a single person. Thus, the required know-how to run and further develop complex, technically sophisticated apps can not be assumed. Therefore, the staff of the library has to be trained adequate. The staff has to support the technical progress.

Libraries using AR have to be aware of overstimulation. Secondary issues like liability and safety have to be taken into account.

5) Considerations of Users

The interviewees where asked for their opinion whether users of libraries (i.e., readers) might have doubts when libraries offer AR apps. First of all privacy aspects have been mentioned. Users might fear that libraries use (tracking) data for personalization and services or even sell the data to third-party companies.

Five interviewees (members of all groups) take the position that using the AR app must not be compulsive. AR has to be an additional, but optional service that users might utilize if they like – and if they identify a benefit. Another group of five interviewees (2 AR experts, 3 librarians of scientific libraries) hold the opinion that users might perceive AR as disconcerting and not being a serious application. Furthermore, users may not be excluded because their mobile devices do not fulfill the required technological standards.

C. AR Applications for Libraries

The central question of the interviews is related to the relevance of AR applications for libraries [17]. There is no clear opinion whether there is a difference between public libraries and scientific libraries. However, several statements indicate that AR apps are relevant only for larger libraries. The experts where asked whether they consider the following applications to be useful and relevant for libraries:

- Augmented books
- Guided tours
- Searching for Media / Additional Information
- Gamification
- Shelf Maintenance

Table I presents an overview of their opinion [17].

1) Augmented Books

There already exist some examples of augmented book for children's books, reference books, and educational textbooks. Some libraries already offer augmented books. But even those libraries that do not offer augmented books yet, consider them as relevant. Several interviewees argue that it does not make sense to augment all books, but it is appropriate to augment some special books to provide added value.

2) Guided Tours

Six (out of eight) members of libraries consider augmented guided tours in the library or the building relevant. The AR app can provide virtual support while exploring the library which might simplify the utilization of the library.

However, three of the AR experts and two librarians did not expect guided tours as practical applications – above all due to the cost-benefit ratio. Especially for small libraries this kind of application is not practical. Classic, non-augmented tours shall not be eliminated since they provide personal contact.

TABLE I. AR APPS FOR LIBRARIES

Int.	**AR Apps**				
	Augmented Books	***Guided Tours***	***Searching for Media / Additional Info***	***Gamifica-tion***	***Shelf Mainten-ance***
AR1	✓	✗	✓	Tend to ✗	✓
AR2	✓	✗	✓	Tend to ✗	✓
AR3	✓	✗	✓	Tend to ✗	✓
AR4	✓	✓	✓	Tend to ✗	Tend to ✓
PL1	✓	✓	✓	✓	✓
PL2	Neutral	✗	✓	✓	✓
PL3	✓	✓	✓	Tend to ✗	✓
PL4	Neutral	✓	✓	Tend to ✗	✗
SL1	✓	✗	✗	Tend to ✗	✗
SL2	✓	✓	✓	Tend to ✗	✗
SL3	✓	✓	✓	Tend to ✗	✓
SL4	✓	✓	✓	Tend to ✗	✓

Scale: Yes ✓ – Tend to Yes ✓ – Neutral – Tend to No ✗ – No ✗
Interviewees (Int.): AR = experts in augmented reality, PL = librarians working in public libraries; SL = librarians working in scientific libraries

3) Searching for Media / Additional Information

All except one interviewee think that searching for media in the library using AR is highly relevant and provides added value to users. Guiding the way to the book in the bookshelf and navigating the user with an AR app on a private mobile device is considered to be useful, especially for large libraries.

However, much more important – according to the interviewees – is the added value that is provided by additional information that can be displayed on the AR device. Additional information especially includes similar media, reviews, and ratings. Searching for a book might result in a list of, for example, four books. Then the AR app fades out all the other books and the four books of interest can be easily identified in the bookshelf. By providing additional information, like the abstract, a summary, or the first chapter of the book, users are assisted in their decision whether a book shall be lend.

Some interviewees (2 AR experts, 3 librarians of scientific libraries) mentioned the advantage of linking online and offline holdings. In addition to physical books in the bookshelf the AR app might display other books that are not available because another user has borrowed it or similar books that might be relevant for the lecture as well. Relevant media might be downloaded to the personal virtual library.

Three librarians argued that a disadvantage of this kind of AR app is its complexity and the amount of work on maintaining this app. The app has to be integrated with several databases in real time to access the additional information needed. The AR experts suggest that using this kind of app on a smartphone will be unsuitable and better for data goggles.

4) Gamification

The usefulness of gamification in the library using augmented reality is considered to be not very high – but it depends on the target group of users. Due to the profile of the users of scientific libraries gamification is hardly relevant. However, if a young audience shall be addressed by a public library AR-based gamification might be an appropriate approach with high potential. Visiting a library and exploring its offerings can be made more thrilling and exciting for young people with the help of games on a mobile devices. An integrated approach based on AR offers several opportunities – for fun, but also to communicate how a library works.

5) Shelf Maintenance

Support on maintaining the bookshelfs of a library using an AR app (like ShelvAR, see Section III.B) would be a real benefit for librarians. Most interviewees (4 AR experts and 5 librarians) agree on that issue. The biggest benefit is the saving of labor and making work much more easier in daily business. Three librarians (one working in a public library, two working in a scientific library) do not expect that an AR app might save much working time.

Shelf maintenance could be done using Internet of Things technologies as well, especially with RFID antenna [21] [22] [23]. Since approaches based on RFID are prone to errors and the implementation is expensive, it is seldom used. Therefore, AR-based approaches might be useful.

D. Findings

The central research question concerns the fact which AR applications are relevant for users of a library. The highest potential is awarded to searching for media and navigating to the correct location in the bookshelf. Searching for media and the display of ancillary information, like ratings, reviews, secondary media, links, etc., is most helpful for users. Using filtering and selection only relevant information can be displayed [17].

Maintaining the bookshelfs of a library with the help of an AR app is a real benefit for librarians. However, this requires a complex implementation and a complex infrastructure. Librarians are interested in providing guided tours through libraries using AR-based, custom designed apps.

Augmented books are an interesting offer for users, but providing augmented books is more in the scope of publishing houses than in the scope of the library. Gamification is the least relevant topic since is requires a very special setup and it is designed only for children and very young people.

Libraries offering AR apps to their users will benefit from improved image, improved quality of service for users, increasing time efficiency and making work easier for librarians.

The biggest challenges when implementing services based on augmented reality apps are cost-benefit ratio, technical issues (like display type, e.g., smartphone vs. data goggles; error rate of tracking), external and internal factors, overstimulation, and doubts of users.

V. Prototype of an AR App for a Scientific Library

To demonstrate the feasibility of an augmented reality app for libraries in practice a prototype app was developed for a university library [24]. The aim of this prototype is to ease the utilization of the library for students and university lecturers and to motivate students to use the library more often by providing additional benefits. The app shall be accessed by smartphones or tablet computer owned by the students. The *Augmented Reality Toolkit Wikitude* by Wikitude is used as the software development kit [25].

Following the fields of application demonstrated in section III two scenarios have been chosen:

- Presentation of additional information at special points of interest
- Display of additional context information to books

Visual triggers are used to activate the display of additional information [24].

A. Augmentations at Points of Interest

Augmentations are triggered by image markers attached to points of interest in the library. The markers consist of a common part identical to all markers (logo, icon) and an individual part including text plus custom icons that can be easily identified by the tracking software (Figure 4).

Fig. 4. Visual marker [24]

Each of the markers is assigned special virtual objects:

- *Welcome* (“Willkommen”) triggers a short video on the library. It is attached at the entrance of the library.
- *Library rules* (“Bibliotheksordnung”) will display the library rules as a PDF document. The markes is attached near the entrance and the library desk.
- *Operating instructions* (“Bedienungsanleitung”) triggers the display of step-by-step instructions on how to use the book scanner. Thus, it is attached on the scanner.

Typically, virtual objects are displayed when the camera of the mobile device is aligned to the marker. Since the information mentioned above is quite complex and requires some time to read, the information is presented in an extra browser window that remains open even if the marker is not identified any more (Figure 5).

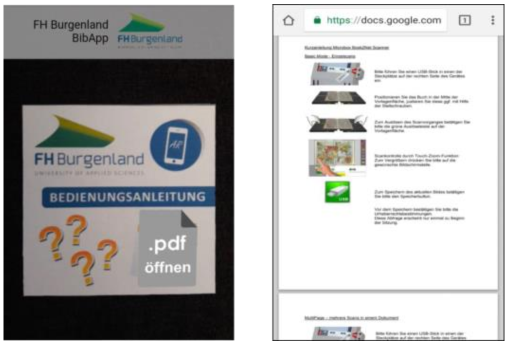


Fig. 5. Operating instructions [24]

B. Context Information for Books

Additional context-based information shall be displayed when the camera of the mobile device points at a specific book. Trigger may be either the book signature or the front cover of the book. The ancillary information is derived from *OPAC* (Online Public Access Catalogue) and the service *Syndetics Unbound* (e.g., summary, information on the author, reviews, ratings).

If the book of interest is in the bookshelf only the spine of the book is visible to the camera. The signature of the book is attached to the spine and can be used to identify the book. According to the signature, ancillary information like full title, names of all authors, language, year of publication, availability as ebook, summary, information on the author(s) will be displayed on the mobile device.

If the front cover of the book is visible, most of the relevant information is already visible and does not have to be augmented. Thus, the following information will be presented: language, year of publication, topic, availability as ebook, summary, information on the author(s) (Figure 6). Due to the length of the text of the summary and the information on the author(s) both types of information are displayed in an additional browser window. The size of the cover pictures used as markers in Wikitude should be minimum 500x500 pixels and maximum 1000x1000 pixels.

Fig. 6. Context information on cover trigger [24]

Because only a few reviews and ratings are available in the database this kind of information is not displayed in the prototype, yet.

C. Features

The prototype offers the following features:

- Recognition of specially designed visual markers for tracking
- Recognition of book covers for tracking
- Retrieval of metadata to identified books (cover marker) from the Wikitude cloud
- Display of text information related to books
- Display of virtual buttons on the touchscreen (Figure 6)
- Play a video
- Display of PDF documents

However, there are still some limitations. Character recognition of the book signature does not work properly in Wikitude. However, some plugins are available yet that might be used to integrate third-party text recognition modules. Since Wikitude works well on image recognition and not on text recognition, book covers that are monochrome and include only text can be hardly recognized. Book covers with high similarity (e.g., a series of books) can not be distinguished as well. They would have to be identified by their signature. Wikitude does not provide a database integration. Thus database access has to be implemented using an additional web service.

VI. Conclusion

A study based on qualitative interviews with librarians working in public and scientific libraries in Austria and experts in augmented reality derives which kinds of applications have potential for libraries. Searching for media, navigating to the correct location and displaying ancillary information, like ratings, reviews, secondary media, links, etc., has the highest potential for users according to the interviewees. AR apps for maintaining the bookshelfs of a library provide real benefit for librarians and are awarded high potential as well. Guided tours through libraries using AR-based, custom designed apps lightens the load of the librarians to introduce new users to the library. Augmented books and gamification approaches are awarded much less potential for libraries. This study provides useful insights on the relevant fields of application of an AR app.

Those insights have been the starting point for the development of a prototype app. It was designed for a university library to demonstrate how an AR app supporting readers (casual users and regular users of a library as well) might look like. The main focus of the prototype is on providing additional information to books that can be identified using visual tracking. The prototype has been designed to identify critical aspects and problem areas when implementing such a service.

VII. Future Work

The survey provided sufficient baseline information to start the prototype implementation of the AR app. However, we interviewed experts in the field (AR experts and librarians). We still have determine the requirements of readers who visit the library and will use this app.

The prototype has limited functionality that has to be enhanced in the future. More databases have to be integrated to provide real time data on additional information related to books. Visual tracking of the signature of books has to be improved to unambiguously identify each book in the library. Books in the bookshelf only show their spine including the signature as a text code. Our prototype does not provide proper character recognition yet which will have to be added in the future. Guiding users to the bookshelf where their book of interest is located has not been tackled yet since this a complex task to achieve in an indoor environment.

References

[1] D. Schmalstieg, and T. Höllerer, Augmented Reality. Boston: Addison Wesley, 2016.

[2] W. Broll, "Augmentierte Realität" in R. Dörner, W. Broll, P. Grimm, B. Jung (Eds.), Virtual und Augmented Reality (VR/AR). Berlin: Springer Verlag, 2013, pp. 241-294.

[3] K. Panetta, 5 Trends Emerge in the Gartner Hype Cycle for Emerging Technologies 2018. Gartner, August 16, 2018 https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/

[4] L. Freyberg, and S. Wolf, "Dienstleistungen einer SmART Library – Anwendungspotentiale von Augmented Reality in Bibliotheken". Medienproduktion - Online-Zeitschrift für Wissenschaft und Praxis, No. 9, pp. 11-15, 2016.

[5] P. Siddappa, librARi. 2014. http://www.pradeepsiddappa.com/design/librari/.

[6] S. Wolf, and S. Büttner. "Mobile Anwendungen in Bibliotheken". Bibliotheksdienst, Vol. 49, No. 1, pp. 14–21, 2015.

[7] K. Ceynowa, "Information „On the Go“: Innovative Nutzungsszenarien für digitale Inhalte – Die Augmented-Reality-App „Ludwig II.“ der Bayerischen Staatsbibliothek". Bibliothek Forschung und Praxis, Vol. 36, pp. 64-69, 2012.

[8] R. Azuma, "A Survey of Augmented Reality". Presence: Teleoperators and Virtual Environments, vol. 6, no. 4, pp. 355-385, August 1997.

[9] P. Milgram, and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays". IEICE Transactions on Information and Systems, vol. E77-D, no. 12, pp. 1321-1329, 1994.

[10] P. Milgram, H. Takemura, A. Utsumi, F. and Kishino, "Augmented Reality: A class of displays on the reality–virtuality continuum". SPIE Vol. 2351, Telemanipulator and Telepresence Technologies, pp. 282-292, 1994.

[11] M. Tönnis, Augmented Reality - Einblicke in die Erweiterte Realität. Berlin: Springer, 2010.

[12] C. Frick, and S. Lange-Mauriège, "Augmented Reality. Anwendungsmöglichkeiten in Bibliotheken". B.I.T. Online, vol. 20, no. 1, pp. 7-14, 2017.

[13] J. Hahn, B. Ryckman, and M. Lux, "Topic Space: Rapid Prototyping a Mobile Augmented Reality Recommendation App". code{4}lib Journal, iss. 30, October 2015. https://journal.code4lib.org/articles/10881

[14] W. Pluta, ShelvAR - Augmented Reality für Bibliotheken. 2011. https://www.golem.de/1104/82994.html.

[15] ShelvAR. http://www.shelvar.com/.

[16] G. Armstrong, J. Hodgson, F. Manista, and M. Ramirez, "The SCARLET Project: Augmented Reality in special collections". SCONUL Focus, no. 54, pp. 52-57, 2012.

[17] B. Baumgartner-Kiradi, "Potential für Augmented Reality Anwendungen für öffentliche und wissenschaftliche Bibliotheken in Hinblick auf das Informationsverhalten der Nutzerinnen und Nutzer". Master Thesis, University of Applied Sciences Burgenland, Austria, 2018.

[18] P. Atteslander, Methoden der empirischen Sozialforschung. Berlin: Erich Schmidt, 13th ed., 2010.

[19] J. Gläser, and G. Laudel, Experteninterviews und qualitative Inhaltsanalyse: Als Instrumente rekonstruierender Untersuchungen. Wiesbaden: VS Verlag für Sozialwissenschaften, 4th ed., 2010.

[20] P. Mayring, Qualitative Inhaltsanalyse: Grundlagen und Techniken. Weinheim: Beltz, 11th ed., 2010.

[21] M. Wójcik, "Internet of Things – potential for libraries". Library Hi Tech, vol. 34, iss. 2, pp. 404-420, 2016. https://doi.org/10.1108/LHT-10-2015-0100.

[22] J. Hahn, "The Internet of Things: Mobile Technology and Location Services in Libraries". Library Technology Reports. Vol. 53, No. 1, 2017.

[23] K. Stefanidis, and G. Tsakonas, "Integration of library services with internet of things technologies". code{4}lib Journal, iss. 30, October 2015. https://journal.code4lib.org/articles/10897.

[24] M. Haberler, "Machbarkeitsstudie einer Augmented Reality App für die Hochschulbibliothek der Fachhochschule Burgenland". Bachelor thesis, University of Applied Sciences Burgenland, Austria, 2018.

[25] Wikitude. https://www.wikitude.com/.

Augmented Reality for Industry 4.0: Architecture and User Experience

Andreas Jakl, Lucas Schöffer, Matthias Husinsky and Markus Wagner
Institute of Creative Media Technologies, St. Pölten University of Applied Sciences, Austria
Email: [first].[last]@fhstp.ac.at

***Abstract*—For *Industry 4.0* – the Internet of Things (IoT) in an industrial manner – new methodologies for support and collaboration of employees are needed. One of these methodologies combines existing work practices with support through technologies like Augmented Reality (AR). Therefore, usability concepts for appropriate hardware as well as the data transfer need to be analyzed and designed within applicable industry standards.**

In this paper, we present two different use cases (*Real-Time Machine Data Overlay* and *Web-Based AR Remote Support*) in the context of collaboration and support of employees. Both use cases are focusing on three main requirements: 1) Effective data transmission; 2) Devices certified for industrial environments; and 3) Usability targeted towards industrial users. Additionally, we present an architecture recommendation for combining both use cases as well as a discussion of the benefits and the limitations of our approaches leading to future directions.

***Index Terms*—Augmented Reality Architecture, Collaboration, Interactive Visualization, Real-Time Data Display, Industry 4.0**

I. Introduction

Many current Augmented Reality use cases target commercial and industrial areas. These have the biggest market potential of the enterprise segment according to the forecast by Goldman Sachs [1]. As such, companies are using pilot projects to evaluate both the real-world usability, as well as the integration into existing work practices.

Typical scenarios involve remote assistance and monitoring. These use cases yield a very measurable return on investment, thereby easing the creation of a business case to offset the research and development costs. Sending specialized engineers to remote locations to fix issues with broken machines is more expensive than developing and providing means for efficient remote collaboration using an on-site generalist, interacting with the remote specialist.

However, the main challenge is designing and creating a suitable software architecture and user interface. Thus, we explicitly focus on the following in this paper:

1) **Efficient data transmission** for mobile scenarios, involving collaboration and / or real-time machine data display.
2) Software should run on durable and cheap **devices** that are **certified for industrial environments**. Most current AR devices are targeted for home or entertainment use.
3) **Usability targeted towards industrial users**, to reduce required training time to a minimum.

To provide guidance for these challenges, we have developed two prototypes: "Web-Based AR Remote Support" and "Real Time Machine Data Overlay".

II. State-of-the-Art for AR in Industry

AR applications for industry got serious research attention since the early 1990's. [2], [3], [4] provided general surveys on AR technologies and frameworks. For industry, AR assistant systems are successfully applied to support humans in training for or during assembly and maintenance processes, as well as quality inspection. By decreasing the mental load, the human error rate decreases while speed is increased at the same time [5], [6], [7], [8].

The advantages of mobile AR for visual analytics for industrial IoT (IIoT) data in a networked environment have been widely recognized [9], [4], [10], [11]. AR technologies provide means to visualize and auralize cyber-physical production systems (e.g., machines and their digital twin [12]) on-site in a context-sensitive way [13], for collaborative robotics [14] and during planning processes [7]. UX design and ergonomics for humans is another topic of great importance for successful implementation in the field [15], [16].

Remote support and collaboration in industrial settings have been discussed for example in [17], [18]. The previous work indicates the demand for AR in industrial scenarios; it presents concepts and solutions for specific issues. However, we found a gap in research concerning the architecture of such solutions – especially while still maintaining a good overall user experience.

III. Our Approaches

Based on the requirements of two Industry 4.0 companies, we have developed two separate prototypes to address their real-life issues. In this section, we first summarize the individual learnings regarding the challenges highlighted in the introduction. Then, we abstract the architecture to provide our recommendations for a generic framework, following best-practices learned through the individual projects.

The first prototype allows placing holographic dashboards in the real world to visualize real-time machine data retrieved through industry standard protocols (*Real-Time Machine Data Overlay*). The second enables a remote expert to draw support annotations on the AR camera view of an on-site generalist (*Web-Based AR Remote Support*).

Fig. 1. Real-time machine data holographic dashboard showing the temperature of a pipe.

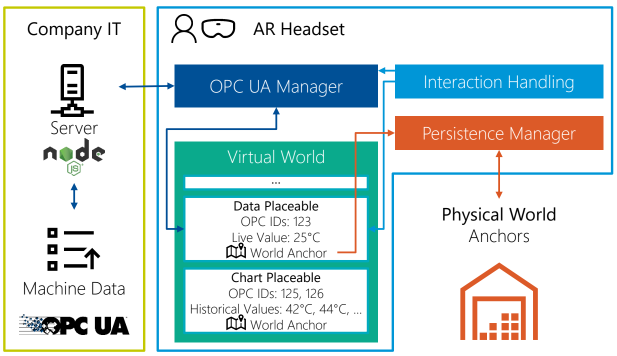


Fig. 2. Possible architecture of an Augmented Reality app, overlaying real-time machine data to actual machines.

A. *Real-Time Machine Data Overlay*

In Fig. 1, a user placed an information dashboard on a pipe. Data is retrieved from an *OPC Unified Architecture* (OPC UA) server. The HoloLens app can show a list of available nodes, whose dashboards can be placed on real-world structures. Interaction is performed with the standardized HoloLens *select* gesture using two fingers. The user interface is mainly composed of *Holographic Buttons* from the MR Toolkit [19]. These are built with semi-transparent elements, which make it easier to retain the view of the real world below the virtual objects. The location of placed items persists across sessions.

The overlay of the most important information directly on the machines in a manufacturing hall allows identifying potential issues with a quick glimpse. The main challenge is combining usability with efficient data transfer (see Fig. 2).

The architecture is based on three separate parts: 1) Company IT containing data storage and processing servers; 2) Mixed Reality (MR) headset handling user interaction and data visualization (VIS); and 3) Actual physical world anchoring of VIS objects.

1) Data Transmission / Company IT: Common interoperability standards for machine data monitoring are OPC UA and *Message Queuing Telemetry Transport* (MQTT).

In the scenario of an untethered MR headset in a factory, data transmission quality and speed can be an issue. As such, a relay server has two advantages: it acts as an additional security layer, as there is no need to expose the OPC UA endpoints to mobile clients in the wireless network. Moreover, it only sends the necessary data to the mobile client to reduce data traffic.

Dürkop et. al [20] analyzed the overheads and data transfer rates of industry protocols with cellular network protocols. Even though the binary variant of the OPC protocol was the most efficient of the analyzed machine-to-machine protocols, it still had a rather large protocol overhead. To reduce this overhead, we added the intermediate Node.js server, as we could strictly isolate the information required by the head mounted display.

2) Devices / AR Headset: The most critical part of the application runs within the actual headset. On the technical side, it needs to register with the server and update its internal database of accessible OPC Node IDs. These are shown in menus, allowing the user to freely place dashboards in the real world. Each dashboard mainly needs to store the connected Node ID and the current dataset in addition to a world anchor.

To be used within industry environments, careful checks need to be performed if the headset fulfills necessary safety standards. The HoloLens already complies to several standards: ANSI Z87.1, CSA Z94.3 and EN 166 [21].

3) Usability / Physical World & Persistence: A key consideration of AR in industrial scenarios is the structure of the real-world environment. The underlying computer vision algorithms from Google ARCore and Microsoft HoloLens use a Simultaneous Localization and Mapping (SLAM) based approach to create a geometrical reconstruction of the world while at the same time estimating camera localization. Even though the commercial algorithms are not available, a state-of-the-art open source implementation is ORB-SLAM [22]. This is a good reference for understanding the underlying algorithms and helped us optimizing the use cases. One of the main steps is detecting key-points in the live camera image. Established approaches detect corners based on the contrast in circular surroundings of individual pixels [23].

Based on this limitation, AR applications should not encourage placing dashboards on feature-less walls (e.g., single-colored with little structure). A better approach is directly placing the items on machines, typically having a more complex structure. In addition, early research from Boeing [5] recommends placing instructions close to the work area.

The world anchor is managed through the Microsoft MR toolkit and essentially forms the connection of the virtual objects to the physical world. Nevertheless, a *persistence manager* component within the app needs to ensure the persistence across sessions, and potentially also across multiple users simultaneously viewing the same scene with different headsets. To avoid instability, virtual objects are recommended to be placed at a maximum distance of $3\,\mathrm{m}$ from an anchor [24].

Special attention needs to be paid to the legibility of the dashboards. Research shows that diegetic and spatial user

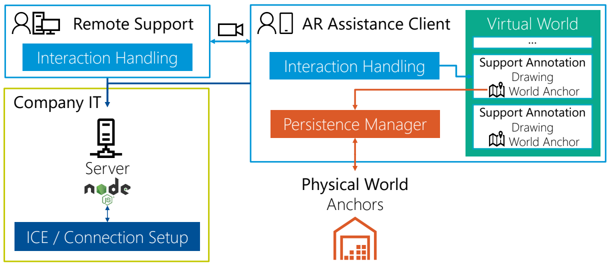


Fig. 3. Possible architecture of a web-based AR support system.

interfaces are the most natural and preferred metaphors for virtual scenarios [25]. However, a limitation of many of today's headsets is the rather low display resolution (HoloLens: 720p [26]). With an ideal placement of dashboards $2\,\mathrm{m}$ away from the user (as recommended by Microsoft guidelines to ensure good focus [27]), small text objects easily become difficult to read. Thus, the dimension of a dashboard showing a numeric live value is $0.2\,\mathrm{m}$ by $0.1\,\mathrm{m}$ in our prototype.

To put usability first, billboard functionality is used for user oriented dashboard visualization. The downside is that parts of the dashboard might appear to be "inside" the physical object. However, due to the semi-transparent material and the deactivated occlusion with the spatial map, that effect is hardly noticeable in our prototype. To optimize anchoring the virtual dashboard on the real-world object, the prototype additionally allows re-positioning through hand-based drag gestures. Manual label placement allows contextual relevance for users and avoids challenges of automated placement [28].

B. Web-Based AR Remote Support

The purpose of remote support via AR is that two persons – a *customer* who needs help on a certain technical problem and a remote expert (*supporter*) who provides a solution on that issue – can exchange visual information (see Fig. 3).

The initial starting point is a video stream of the customer's environment. The customer shows the point of interest (POI, describes the object where the actual problem is located) to the supporter. In industrial environments this could be a malfunctioning machine. AR technology allows the supporter to give the customer interactive visual feedback (e.g., by making annotations), which is anchored to the POI and stays in place.

For our system we identified two core requirements during the conceptual phase: first, the supporter should have no need to install additional software. Second, no special hardware should be required – neither from the customer nor from the supporter. We decided to use mobile devices (smartphones or tablets) for the customer, as these are widespread and available in industry-certified variants that can be used in environments like production halls.

Overall, the system should be universally applicable, as well as easily available for a huge potential target group.

1) Devices & Installation: In our scenario, the customer is willing to install new software on his mobile device, if it helps him to get quick and efficient support. However, the effort for that installation has to be low. After installation, the app should be usable without the need for configuration.

The supporter was defined as an expert with special knowledge on certain machines or technical devices. He maybe travels a lot (e.g., doing installations). There's a chance that he has no possibility or time to install software when the customer calls for quick help.

To satisfy these requirements, we designed a system where the supporter gets an individual hyperlink (via E-Mail) from the customer. After opening the link, the supporter receives the real-time video stream from the customer within his web browser. There, he can add visual feedback, environmentally linked to the Augmented Reality POI.

2) Data Transmission: The underlying technology for the data transmission of the stream and the graphical annotations is WebRTC. WebRTC is a web standard for building peer-to-peer connections between two browsers or between a browser and another application that supports an implementation of WebRTC. During the stream, there is no need for an additional node or logical overhead in-between (e.g., a streaming server).

To establish the connection, the customer's app automatically registers with a signaling server and receives a session ID. The app then generates the individual hyperlink, which includes the new session ID. The customer sends that hyperlink to the supporter (e.g., via E-Mail). By opening the hyperlink, the supporter's web browser also connects to the signaling server. Thus, the two peers exchange their Interactive Connectivity Establishment (ICE) information via the signaling server and establish a direct peer-to-peer connection. Next, the customer streams the environment to the supporter by using the rear camera of his mobile device. At the same time, the device performs environmental understanding on the captured video input using ARCore, attempting to find key-points.

3) Usability & Collaboration: For drawing on a video stream, there are two different concepts: (a) The drawings can be integrated in real-time into the currently active stream. This could cause inaccuracies when the customer moves the device while the supporter is drawing. (b) Pause the video stream in the supporter's browser during drawing.

Based on these two general approaches, the focus of the project on accurate annotations and usage within industry environments leads to approach (b): whenever the supporter performs a tap on the streamed video, the last frame freezes in the browser. The supporter then has the possibility to draw annotations on the frozen frame. At the same time, the customer can continue to move his phone without influencing the supporter's view.

At the same time, a transparent plane is created as AR element in the environment of the customer. By setting an anchor (using the detected environmental key-points), the plane stays at the designated location near the POI. The annotation will later appear on that plane.

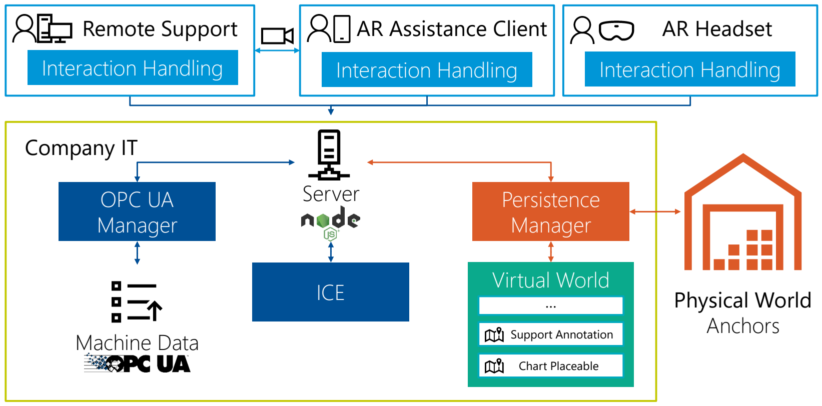


Fig. 4. Possible architecture of an combined approach.

On completion, the supporter confirms his annotations. The annotation data is directly sent to the customer. There, it is used to create an image texture, which is added to the previously generated plane. The supporter's annotation becomes visible on the plane and is anchored to the POI. The annotations appear in the customer's environment.

IV. Recommendations for AR App Architecture

These two individual use cases are tightly related to each other. Combined, they lead to a complete use case, while still retaining the unique architectural challenges of each part. While our approach A describes the initial dashboard view for live machine data, approach B then allows connecting to a remote expert / supporter in case issues become evident.

In Fig. 4 we show an architecture unifying both use cases. It exhibits two main differences: 1) The client's responsibility is focused only on interaction handling – allowing the users to add new dashboards, or to draw annotations; 2) The persistence is centralized in the server. This creates a shared database, allowing improved multi-user support, as well as a more seamless transition between both use cases. For example, improvements suggested by the supporter could be immediately seen in the real-time machine data dashboards.

This centralized architecture is based on the most recent developments to share anchors between users and platforms. While SLAM-algorithms from researchers have already been optimized for collaborative SLAM [29], commercial implementations are currently also adding support for environment data sharing between multiple clients (Google Cloud Anchors [30], Apple Shared Experiences [31]). These new APIs will be a key enabler for future improvements to our system.

Based on the initial observations from our prototypes, we see the distinction between different specialized clients as an important factor. Head-mounted displays generally have a higher cost and might be less comfortable to wear for a whole working day, but offer better immersion and precision (e.g., through time-of-flight depth sensing). Such a device would be suitable for example for the shift manager to get a quick overview of dashboards of various machines. On the other hand, smartphone-based AR works on industry-ready phones, which are easier to carry and cheaper to roll out to employees in service and production. Overall, implementing a common shared persistence back-end with specialized clients for the different scenarios gives the best of both worlds.

V. Conclusion and Future Directions

We presented two different use cases (see Sec. III) in the context of collaboration and support of employees (*Real-Time Machine Data Overlay* and for *Web-Based AR Remote Support*), meeting the future challenges of *Industry 4.0*. Based on the two approaches and considering our lessons learned, we proposed a future design approach for a combined architecture (see Sec. IV).

Since our prototypes have only been evaluated at a very low level (non-documented discussions with domain experts), it is necessary to perform a design study [32] and evaluation [33] for the proposed architecture. Therefore, this approach needs to be designed and developed in a user-centered process [34] where future system users are fully included in the evaluation cycle. Additionally, a usability study is needed to evaluate the integrated visualizations as well as the general workflow concept regarding industry employees.

The insights generated through our two prototypes, combined with previously gathered experience of collaborative and multi-device scenarios [35], will provide a profound base for these planned further research activities. However, to cover all the new upcoming challenges for *Industry 4.0* more research is needed involving the employees directly into the design and conception loop to not get overwhelming by the new created technologies.

Acknowledgment

The authors would like to thank the FFG and the companies for supporting these projects.

References

[1] H. Bellini, W. Chen, M. Sugiyama, M. Shin, S. Alam, and D. Takayama, "Virtual & Augmented Reality: Understanding the race for the next computing platform," *Profiles in Innovation*, pp. 1–30, Jan. 2016. [Online]. Available: https://www.goldmansachs.com/insights/pages/technology-driving-innovation-folder/virtual-and-augmented-reality/report.pdf

[2] D. Van Krevelen and R. Poelman, "A survey of augmented reality technologies, applications and limitations," *International journal of virtual reality*, vol. 9, no. 2, pp. 1–20, Jun. 2010.

[3] M. Billinghurst, A. Clark, and G. Lee, "A Survey of Augmented Reality," *Foundations and Trends® in Human–Computer Interaction*, vol. 8, no. 2-3, pp. 73–272, Mar. 2015. [Online]. Available: http://www.nowpublishers.com/article/Details/HCI-049

[4] D. Chatzopoulos, C. Bermejo, Z. Huang, and P. Hui, "Mobile Augmented Reality Survey: From Where We Are to Where We Go," *IEEE Access*, vol. 5, pp. 6917–6950, 2017.

[5] T. P. Caudell and D. W. Mizell, "Augmented reality: an application of heads-up display technology to manual manufacturing processes," in *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*, vol. ii, Jan. 1992, pp. 659–669 vol.2.

[6] Z. Zhu, V. Branzoi, M. Wolverton, G. Murray, N. Vitovitch, L. Yarnall, G. Acharya, S. Samarasekera, and R. Kumar, "AR-mentor: Augmented reality based mentoring system," in *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Sep. 2014, pp. 17–22.

[7] U. Bockholt, S. Brauns, O. Fluck, A. Harth, P. Keitler, D. Koriath, S. Lengowski, M. Olbrich, I. Staack, U. Rautenberg, and V. Widor, "Werkerassistenz," in *Web-basierte Anwendungen Virtueller Techniken*. Springer Vieweg, Berlin, Heidelberg, 2017, pp. 309–377. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-662-52956-0_7

[8] G. Dini and M. D. Mura, "Application of Augmented Reality Techniques in Through-life Engineering Services," *Procedia CIRP*, vol. 38, pp. 14–23, Jan. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2212827115008033

[9] E. Siow, T. Tiropanis, and W. Hall, "Analytics for the Internet of Things: A Survey," *arXiv:1807.00971 [cs]*, Jul. 2018, arXiv: 1807.00971. [Online]. Available: http://arxiv.org/abs/1807.00971

[10] A. Stork, "Visual Computing Challenges of Advanced Manufacturing and Industrie 4.0 [Guest editors' introduction]," *IEEE Computer Graphics and Applications*, vol. 35, no. 2, pp. 21–25, Mar. 2015.

[11] J. Posada, C. Toro, I. Barandiaran, D. Oyarzun, D. Stricker, R. d. Amicis, E. B. Pinto, P. Eisert, J. Döllner, and I. Vallarino, "Visual Computing as a Key Enabling Technology for Industrie 4.0 and Industrial Internet," *IEEE Computer Graphics and Applications*, vol. 35, no. 2, pp. 26–40, Mar. 2015.

[12] M. Grieves, "Digital Twin: Manufacturing Excellence through Virtual Factory Replication," Florida Institute of Technology, Tech. Rep., 2014. [Online]. Available: http://innovate.fit.edu/plm/documents/doc_mgr/912/1411.0_Digital_Twin_White_Paper_Dr_Grieves.pdf

[13] acatech – Deutsche Akademie der Technikwissenschaften, Ed., *Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry. Final report of the Industrie 4.0 Working Group*, Apr. 2013. [Online]. Available: http://publica.fraunhofer.de/dokumente/N-275303.html

[14] S. A. Green, M. Billinghurst, X. Chen, and J. G. Chase, "Human-Robot Collaboration: A Literature Review and Augmented Reality Approach in Design," *International Journal of Advanced Robotic Systems*, vol. 5, no. 1, p. 1, Mar. 2008. [Online]. Available: https://doi.org/10.5772/5664

[15] D. Bowman, E. Kruijff, J. J. L. Jr, and I. P. Poupyrev, *3D User Interfaces: Theory and Practice, CourseSmart eTextbook*. Addison-Wesley, Jul. 2004, google-Books-ID: JYzmCkf7yNcC.

[16] M. Peissner and C. Hipp, *Potenziale der Mensch-Technik Interaktion für die effiziente und vernetzte Produktion von morgen*, D. Spath and A. Weisbecker, Eds. Fraunhofer Verlag.

[17] S. Bottecchia, J.-M. Cieutat, and J.-P. Jessel, "T.A.C: augmented reality system for collaborative tele-assistance in the field of maintenance through internet," in *Proceedings of the 1st Augmented Human International Conference on - AH '10*. Megève, France: ACM Press, 2010, pp. 1–7. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1785455.1785469

[18] L. Alem and W. Huang, *Recent Trends of Mobile Collaborative Augmented Reality Systems*. Springer Science & Business Media, Sep. 2011, google-Books-ID: HGxZytjBBnsC.

[19] "Interactable Objects & Receiver," Nov. 2018, accessed: Sept. 19^{th}, 2018. [Online]. Available: https://github.com/Microsoft/MixedRealityToolkit-Unity/blob/120e52afac62cc3c1a2948ccf59a372c96578473/Assets/HoloToolkit-Examples/UX/Readme/README_InteractableObjectExample.md

[20] A. Neumann, M. J. Mytych, D. Wesemann, L. Wisniewski, and J. Jasperneite, "Approaches for In-vehicle Communication – An Analysis and Outlook," in *Computer Networks*, P. Gaj, A. Kwiecień, and M. Sawicki, Eds. Cham: Springer International Publishing, 2017, vol. 718, pp. 395–411. [Online]. Available: http://link.springer.com/10.1007/978-3-319-59767-6_31

[21] Microsoft, "Microsoft HoloLens," accessed: Sept. 14^{th}, 2018. [Online]. Available: https://www.microsoft.com/en-us/hololens/commercial-overview

[22] R. Mur-Artal and J. D. Tardós, "ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, Oct. 2017.

[23] E. Rosten and T. Drummond, "Machine Learning for High-Speed Corner Detection," in *Computer Vision – ECCV 2006*, ser. Lecture Notes in Computer Science, A. Leonardis, H. Bischof, and A. Pinz, Eds. Springer Berlin Heidelberg, 2006, pp. 430–443.

[24] A. Turner, M. Zeller, E. Cowley, and B. Bray, "Spatial anchors - Mixed Reality." [Online]. Available: https://docs.microsoft.com/en-us/windows/mixed-reality/spatial-anchors

[25] P. Salomoni, C. Prandi, M. Roccetti, L. Casanova, L. Marchetti, and G. Marfia, "Diegetic user interfaces for virtual environments with HMDs: a user experience study with oculus rift," *Journal on Multimodal User Interfaces*, vol. 11, no. 2, pp. 173–184, Jun. 2017. [Online]. Available: http://link.springer.com/10.1007/s12193-016-0236-5

[26] Microsoft, "Rendering – Mixed Reality," accessed: Sept. 16^{th}, 2018. [Online]. Available: https://docs.microsoft.com/en-us/windows/mixed-reality/rendering

[27] Microsoft, "Comfort – Mixed Reality," accessed: Sept. 12^{th}, 2018. [Online]. Available: https://docs.microsoft.com/en-us/windows/mixed-reality/comfort

[28] R. Azuma and C. Furmanski, "Evaluating Label Placement for Augmented Reality View Management," in *Proceedings of the 2Nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ser. ISMAR '03. Washington, DC, USA: IEEE Computer Society, 2003, p. 66. [Online]. Available: http://dl.acm.org/citation.cfm?id=946248.946790

[29] J. Huai, G. Jóźków, D. Grejner-Brzezinska, and C. Toth, "Collaborative Monocular SLAM with Crowd-Sourced Data," in *Proceedings of the 29th International Technical Meeting of The Satellite Division of the Institute of Navigation*, Portland, Oregon, Sep. 2016, pp. pp. 1064 – 1079.

[30] Google Developers, "Share AR Experiences with Cloud Anchors | ARCore," accessed: Sept. 19^{th}, 2018. [Online]. Available: https://developers.google.com/ar/develop/java/cloud-anchors/overview-android

[31] Apple, "Creating a Multiuser AR Experience | Apple Developer Documentation," accessed: Sept. 19^{th}, 2018. [Online]. Available: https://developer.apple.com/documentation/arkit/creating_a_multiuser_ar_experience

[32] M. Sedlmair, M. Meyer, and T. Munzner, "Design study methodology: Reflections from the trenches and the stacks," *IEEE Trans. Vis. and Comp. Graphics*, vol. 18, no. 12, pp. 2431–2440, Dec. 2012.

[33] K. Blumenstein, C. Niederer, M. Wagner, G. Schmiedl, A. Rind, and W. Aigner, "Evaluating information visualization on mobile devices: Gaps and challenges in the empirical evaluation design space," in *Proceedings of 2016 Workshop on Beyond Time And Errors: Novel Evaluation Methods For Visualization*, ACM. Baltimore, MD, USA: ACM, 2016, pp. 125–132.

[34] H. Sharp, Y. Rogers, and J. Preece, *Interaction Design: Beyond Human-Computer Interaction*, 2nd ed. Chichester; Hoboken, NJ: John Wiley & Sons, Jan. 2007.

[35] K. Blumenstein, M. Wagner, and W. Aigner, "Cross-platform infovis frameworks for multiple users, screens and devices: Requirements and challenges," in *DEXiS 2015 Workshop on Data Exploration for Interactive Surfaces. Workshop in conjunction with ACM ITS'15*, online. Funchal, Portugal: online, 11/2015 2015, pp. 7–11.

It Pays to Be Lazy: Reusing Force Approximations to Compute Better Graph Layouts Faster

Robert Gove
Two Six Labs, robert.gove@twosixlabs.com

***Abstract*—N-body simulations are common in applications ranging from physics simulations to computing graph layouts. The simulations are slow, but tree-based approximation algorithms like Barnes-Hut or the Fast Multipole Method dramatically improve performance. This paper proposes two new update schedules, uniform and dynamic, to make this type of approximation algorithm even faster by updating the approximation less often. An evaluation of these new schedules on computing graph layouts finds that the schedules typically decrease the running time by 9% to 18% for Barnes-Hut and 88% to 92% for the Fast Multipole Method. An experiment using 4 layout quality metrics on 50 graphs shows that the uniform schedule has similar or better graph layout quality compared to the standard Barnes-Hut or Fast Multipole Method algorithms.**

I. Introduction

Spring-electric algorithms are considered to be conceptually simple methods for computing graph layouts [1]–[5] and they have enjoyed widespread implementation. However, the brute force algorithm requires $O(|V|^2)$ time to compute repulsive forces, where $|V|$ is the number of vertices in a graph $G = (V, E)$. Tree-based approximation methods—e.g. Barnes-Hut (BH), the Fast Multipole Method (FMM), and the Well-Separated Pair Decomposition (WSPD)—reduce this running time to $O(|V| \log |V|)$. Many spring-electric algorithms employ these techniques to improve performance [4]–[16], so improving these techniques' speed can have a wide impact. Reducing the amount of computation can reduce energy consumption on battery powered devices, reduce interruptions to analysts' flow of thought and attention [17], and accelerate the visual analytics sensemaking process.

These approximation algorithms create tree-based data structures from the vertex positions, and the algorithms then use the trees to approximate repulsive forces. Because the spring-electric algorithm iteratively updates vertex positions, the approximation methods reconstruct the tree after each iteration using the new vertex positions. Tree construction runs in $O(|V| \log |V|)$ time, and therefore can be computationally costly, but it is unknown whether it is necessary to construct a new tree after every iteration. Many spring-electric algorithm implementations include a "cooling" parameter that reduces the change in vertex position over time, indicating that calculating new trees provides diminishing improvements to accuracy after each iteration of the algorithm.

This paper presents an evaluation[1] of three alternative schedules for updating tree-based approximations less frequently: Logarithmic, uniform, and dynamic. The evaluation compares these schedules to the standard schedule of reconstructing the tree after every iteration. This paper shows that using a logarithmic, uniform, or dynamic update schedule achieves significantly faster performance compared to the standard update schedule. In addition, the uniform schedule achieves the same or better graph layout quality as the standard schedule.

This paper makes the following contributions: (1) A new dynamic algorithm for deciding when to reconstruct trees used in tree-based approximations such as Barnes-Hut, the Fast Multipole Method, or the Well-Separated Pair Decomposition; (2) a new schedule for reconstructing trees at uniform frequency; (3) a reformulation of the angular resolution (dev) readability metric to make it yield a value in $[0, 1]$; and (4) an evaluation of the logarithmic, uniform, and dynamic update schedules compared to the standard update schedule showing that the new uniform schedule outperforms the other methods.

[1]The materials to reproduce the analysis are available at https://osf.io/re7nx/

II. Background

Spring-electric algorithms belong to the family of force-directed graph layout algorithms. Spring-electric algorithms cast the graph layout problem as an iterative physical simulation, where the algorithm models the graph's vertices similarly to charged particles that repel each other, and it models the graph's edges similarly to springs that define an ideal distance between vertices. This paper is concerned with improving the runtime of the repulsive force calculation.

The *Barnes-Hut* (BH) approximation builds a quadtree of vertex positions, and then considers distant groups of vertices as a single large vertex (see Barnes and Hut [18] and Quigley and Eades [19] for more details). This process of calculating the quadtree runs in $O(|V| \log |V|)$ time and reduces the force calculations to $O(|V| \log |V|)$.

The *Fast Multipole Method* (FMM), like Barnes-Hut, first builds a spatial tree based on the vertex positions in $O(|V| \log |V|)$ time, but then it aggregates nodes in the tree in order to calculate repulsive forces in $O(|V|)$ time [20]–[22].

Recently, Lipp et al [4], [5] proposed using the *Well-Separated Pair Decomposition* [23] (WSPD)—another tree-based approximation algorithm—to compute repulsive forces for graph layout algorithms. Both tree construction and repulsive force calculation run in $O(|V| \log |V|)$ time.

For all three of the above tree-based approximation methods, the graph layout algorithm must reconstruct the tree after each iteration because the vertex positions have changed.

However, is it necessary to calculate a new tree after every iteration? Lipp et al [4], [5] experimented with updating the WSPD whenever $\lfloor 5\log(i)\rfloor$ changes (where i is the current iteration number), instead of after every iteration. They found that this can decrease the number of edge crossings [5] compared to the standard update schedule (i.e. reconstructing the tree after every iteration), but they did not compare the running time of the $\lfloor 5\log(i)\rfloor$ method to the standard method, nor did they test other update schedules such as other multiples of $\log(i)$ or a uniform update schedule. They also suggested using a dynamic algorithm to determine when to update the tree [5], but they did not define an algorithm to accomplish this or evaluate this idea. Furthermore, they did not apply this to the Barnes-Hut approximation or Fast Multipole Method, so it is not clear if it will work with other tree-based approximations, or whether their $\lfloor 5\log(i)\rfloor$ update criteria is optimal.

III. Schedules for Updating Approximations

There are many different methods to determine when to update trees used in tree-based approximation methods such as the Barnes-Hut (BH) approximation, the Fast Multipole Method (FMM), or the Well-Separated Pair Decomposition (WSPD). The standard versions of these algorithms construct a new tree after every iteration, but the tree could be reconstructed less often. This could be determined by an algorithm that defines some sort of schedule of when to update the approximation by reconstructing the tree. In this paper, a *schedule* is a function that returns a boolean value indicating whether or not the tree should be reconstructed. This paper explores three alternative schedules to the standard schedule: logarithmic, uniform, and dynamic.

A. Logarithmic Schedule

Many spring-electric algorithm implementations include a "cooling" parameter that reduces the change in vertex position over time, indicating that constructing new trees may provide diminishing improvements to accuracy after each iteration of the algorithm. In addition, most vertices tend to converge to a final position. This motivates a schedule that constructs a new tree with decreasing frequency at later iterations of the force-directed layout. Lipp et al. [4], [5] proposed a logarithmic schedule where a new tree is constructed if $\lfloor 5\log(i)\rfloor$ changes, where i is the current iteration number. However, they did not experiment with using scalars other than 5, and they did not specify the logarithmic base. Note that any two logarithmic functions with different bases differ by only a constant (i.e. $\log_b(x) = \log_a(x)/\log_a(b)$), so it suffices to use a base 10 logarithm and vary the scalar multiple. For this reason, this paper explores the family of schedules defined by $\lfloor k\log(i)\rfloor$ where k is an integer in $[1, 10]$. For 300 iterations of a graph layout algorithm, this corresponds to constructing a new tree 7, 13, 18, 22, 26, 31, 34, 38, 42, or 45 times for k ranging from 1 to 10.

B. Uniform Schedule

Even though vertex velocity tends to decrease over time, velocity does change at every iteration of a force-directed layout algorithm, and the exact amount of change can be difficult to predict. For this reason, it may be desirable to construct a new tree at uniform intervals. This paper proposes the following uniform update schedule: let u_k be the number of trees constructed using the logarithmic schedule for integer k. Then, for a layout that has n iterations, construct a new tree every n/u_k iterations. For a given value of k, this results in the same number of updates for both the logarithmic and the uniform schedules, and supports direct comparisons between them. The difference between the schedules is that, for a fixed value of k, the uniform schedule has the same number of iterations between each tree construction, whereas the logarithmic schedule constructs more trees at the beginning of the layout and fewer at the end.

C. Dynamic Schedule

During a force-directed layout, vertex velocity may temporarily decrease as the layout enters a local minimum before increasing as the layout escapes the minimum. This, combined with the "cooling" parameter described above, motivates a dynamic approach to deciding when to construct a new tree: if vertex positions are changing rapidly, then a new tree can improve accuracy, but if vertex positions are changing slowly, then constructing a new tree may be computationally costly with little improvement in accuracy. Therefore we would like to construct a new tree only if the old tree is out of date.

Algorithm 1 shows a dynamic algorithm that performs a check to decide whether to construct a new tree or use the old tree to calculate repulsive forces. This check keeps a running sum of the velocities (displacement) of all vertices since the last time a tree was constructed. If this running sum exceeds the previous sum, the new sum is assigned to the previous sum, the running sum is reset to 0, and a new tree is calculated before computing repulsive forces; otherwise, the algorithm decides the old tree is accurate enough and uses it to compute forces on the vertices. In practice, this algorithm constructs new a tree about 10–15 times out of 300 iterations.

Algorithm 1 Dynamic schedule. Initially, $currSum = prevSum = 0$. u_{v_x} and u_{v_y} are u's x, y velocities

```
for each vertex u do
    currSum ← currSum + |u_vx| + |u_vy|
if tree is null or currSum ≥ prevSum then
    prevSum ← currSum
    currSum ← 0
    tree ← newTree()
computeForces()
```

This dynamic algorithm has the benefit of generalizing to any approximation method or type of tree used. For example, an alternative dynamic algorithm might operate on the quadtrees in the Barnes-Hut approximation and check the number of vertices that are no longer in their original quadtree cells. However, such a dynamic algorithm would be restricted to approximation methods that depend on quadtrees, such as the Barnes-Hut approximation, and it would not generalize to other tree-based approximation methods. Another downside to this alternative approach is that if only a few vertices moved out of their cells, but they moved a large distance, this may not exceed the threshold to construct a new quadtree. On the other hand, the dynamic method in Algorithm 1 can determine that a new tree should be constructed in this case where only a few vertices have moved but they moved a large distance.

Early experiments with this algorithm tested small coefficients for updating $currSum$ such as

$$currSum \leftarrow c * currSum + |u_{v_x}| + |u_{v_y}|$$

for $c = 1.01$ or $c = 0.99$. However, even such small coefficients resulted in constructing a new tree way more often than necessary (as in 1.01) or not enough to be useful (as in 0.99). Therefore the algorithm uses a coefficient of 1, which seems to yield good results.

IV. Exploratory Analysis

In order to develop hypotheses to test, this section presents an exploratory analysis of the update schedules. All data collection and analysis was conducted in NodeJS version 9.4.0 on a 2015-model MacBook Pro with a 3.1 GHz Intel Core i7 processor and 16 GB of RAM.

This exploratory analysis uses 50 graphs with 1000 vertices or fewer randomly selected from the KONECT [24] and SuiteSparse [25] graph collections.

This evaluation uses the D3.js framework [9] to compare the four update schedules (logarithmic, uniform, dynamic, and standard). The evaluation uses D3.js's default settings, and adds a central gravitational force with a strength of 0.001. The experiment uses D3.js's default stopping criteria, which is when the "cooling" parameter becomes sufficiently small; by default, this occurs after 300 iterations.

D3's default force-directed algorithm uses the Barnes-Hut (BH) approximation, which has been modified for this evaluation to support the new update schedules. This evaluation also uses a second algorithm, which is a modified version of D3's force-directed algorithm but uses a publicly available implementation of the Fast Multiple Method[2] (FMM). This second force-directed algorithm also supports all four update schedules. Although other tree-based approximation methods are available, they are implemented in other programming languages, and therefore cannot be directly compared to the JavaScript implementations. For this reason, this analysis only uses the aforementioned JavaScript versions of the BH and FMM approximations in order to minimize threats to validity.

This evaluation uses 10 versions each of the logarithmic and uniform update schedules parameterized with k from 1 to 10 as described in Section III-A and Section III-B.

D3.js initializes vertices in a disc-like phyllotaxis arrangement, where vertices at the beginning of the vertex array are at the center and vertices at the end of the vertex array are at the periphery. To minimize any possible effects of the initial positions on the experiment results, this experiment randomly shuffles the vertex array before calculating initial positions and running the spring-electric layout algorithm. This is done 20 times for each graph for each pair of update schedule and approximation method ($4 \times 2 = 8$ schedule-approximation pairs). The experiment then records the median runtime and median readability metrics for each graph and algorithm combination (this experiment uses the median instead of the arithmetic mean to avoid the results being skewed by outliers).

This evaluation uses the edge crossing, edge crossing angle, angular resolution (min), and angular resolution (dev) graph layout readability metrics implemented in greadability.js[3], which are defined below. Other readability metrics exist, such as stress or standard deviation of edge length, but this evaluation avoids these readability metrics because they have known issues [10], [26], [27]. Namely, non-uniform edge lengths are often necessary to achieve good layouts for real-world graphs [27], [28]; preserving shortest-path distances, as measured by stress, may not be ideal for producing good layouts [27]; stress is not defined on graphs with more than one component, which often occur in real-world data; and two layouts that convey a graph's structure equally can have different stress values [26]. In addition, stress and standard deviation of edge length do not have normalized versions that support accurate comparisons between different graphs.

Edge crossings, denoted $\aleph_c$, measures the number of edges that cross, or intersect, in the layout. The metric scales the number of edge crossings, c, by an approximate upper bound so that $\aleph_c \in [0, 1]$.

$$\aleph_c = 1 - c \Big/ \left(\frac{|E|(|E|-1)}{2} - \frac{1}{2} \sum_{u \in V} \deg(u)(\deg(u)-1) \right)$$

If the denominator is 0, then $\aleph_c = 1$.

Edge crossing angle, denoted $\aleph_{ca}$, measures the average deviation of each edge's crossing angle from the ideal angle ϑ of 70 degrees.

$$\aleph_{ca} = 1 - \frac{\sum_{e \in E} \sum_{e' \in c(e)} |\vartheta - \theta_{e,e'}|}{c\vartheta}$$

where $c(e)$ is the set of edges that intersect e, and $\theta_{e,e'}$ is the acute angle of the two intersecting edges. If $c\vartheta = 0$, then $\aleph_c = 1$.

Angular resolution (min) is the average deviation of incident edge angles from the ideal minimum angle for each vertex u.

$$\aleph_{rm} = 1 - \frac{1}{|V|} \sum_{u \in V} \frac{|\vartheta_u - \theta_{u_{min}}|}{\vartheta_u}$$

Here, $\vartheta_u = 360/d(u)$, the degrees between each incident edge if the angles were uniform, and $\theta_{u_{min}}$ is the smallest measured angle between edges incident on u.

Angular resolution (dev) is the average deviation of angles between incident edges on each vertex u. This paper presents a new formulation of the angular resolution (dev) metric described by Dunne *et al* [29]. Dunne *et al*'s metric can

[2]https://github.com/davidson16807/fast-multipole-method

[3]https://github.com/rpgove/greadability

produce negative numbers, but the version presented here produces a value in $[0, 1]$.

$$\aleph_{rd} = 1 - \frac{1}{|V|} \sum_{u \in V,\, d(u)>1} \left(\frac{1}{2d(u)-2} \sum_{i}^{d(u)} \frac{|\vartheta_u - \theta_{i,(i+1)}|}{\vartheta_u} \right)$$

Here, $\theta_{i,(i+1)}$ is the acute angle between adjacent edges i and $i+1$ that are incident on vertex u, modulo $d(u)$, the degree of u. To see that $\aleph_{rd} \in [0, 1]$, consider that the maximum deviation on a vertex u will occur when $\theta_{i,(i+1)} \approx 0$ for all incident edges except one where $\theta_{i,(i+1)} \approx 360$. For the central vertex in a star graph with $|V|$ vertices, the sum of the deviations of its incident edges would be

$$\sum_{i}^{d(u)} \frac{|\vartheta_u - \theta_{i,(i+1)}|}{\vartheta_u} = \sum_{i}^{d(u)} \frac{|360/d(u) - \theta_{i,(i+1)}|}{360/d(u)}$$

Since $\theta_{i,(i+1)} \approx 0$ for all but one pair of incident edges, we then have

$$\begin{aligned} \sum_{i}^{d(u)} \frac{|360/d(u) - \theta_{i,(i+1)}|}{360/d(u)} &\approx (d(u)-1) + \frac{|360/d(u) - 360|}{360/d(u)} \\ &= (d(u)-1) + (d(u)-1) \\ &= 2d(u) - 2 \end{aligned}$$

This is the upper bound for the deviation of a vertex u, so we must divide the deviation of each vertex by this quantity. In contrast, the smallest value would occur in a graph where every pair of incident edges had the ideal angle, which would make the deviation 0. Therefore, $\aleph_{rd}$ is in the range $[0, 1]$.

See Dunne *et al* [29] for a more detailed discussion of these and other readability metrics.

A. Hypothesis Generation

Figure 1 shows the median runtime and readability metrics across all 50 graphs in the exploratory dataset. The runtime and readability metrics were calculated for the Barnes-Hut (BH) and Fast Multipole Method (FMM) approximation algorithms using each update schedule (standard, dynamic, logarithmic for k from 1 to 10, and uniform for k from 1 to 10). This gives us some insight into how each approximation algorithm performs for the different update schedules, allowing us to develop hypotheses to formally test.

In the exploratory data, the standard BH and FMM methods are slower than all of the dynamic, log, and uniform update schedules. The dynamic algorithm had about the same number of updates as $k = 2$ (median is 11 and 14 for BH and FMM respectively), and they also have about the same runtime. This leads to **Hypothesis 1**: Using a dynamic or fixed update schedule (e.g. logarithmic or uniform) is faster than the standard update schedule for both the BH and FMM algorithms.

In order to test Hypothesis 1 experimentally, we must choose a value of k to use for the logarithmic and uniform update schedules. Ideally, k will be small in order to minimize the runtime, but k should also be large enough to produce good quality layouts. For these reasons, let us choose $k = 4$ (i.e. constructing a new tree 22 times out of 300 iterations). Although the runtime is not as short as $k = 1$, the layout quality appears substantially better for the majority of update schedules and approximation algorithms. Higher values of k do not appear to provide much improvement in layout quality, and in fact large values of k sometimes produce lower quality layouts (e.g. for number of crossings when $k = 9$ or 10).

For both the BH and FMM algorithms, the dynamic schedule appears to have faster runtime than the logarithmic and uniform schedules for $k > 2$. Therefore, **Hypothesis 2** is that the dynamic schedule will decrease the BH and FMM runtime more than the logarithmic or uniform schedules if $k = 4$.

For BH, the dynamic schedule appears to generally have worse readability metric performance than the standard schedule. **Hypothesis 3** is that the dynamic schedule will perform worse on all readability metrics than the standard schedule, except for edge crossings where the dynamic schedule will perform about the same.

On the other hand, for BH with the logarithmic and uniform schedules where $k = 4$, these schedules tend to have better readability metrics. **Hypothesis 4** is that, for BH, the logarithmic and uniform schedules will have better readability metrics compared to the standard schedule, except for crossing angle where logarithmic and uniform will perform worse.

For FMM, we see similar, but somewhat different, relationships in the readability metrics. By looking at the data, we believe that all other schedules perform better than the standard schedule on the edge crossings and angular resolution (min) metrics (**Hypothesis 5**), all schedules perform worse than the standard schedule on the crossing angle metric (**Hypothesis 6**), and that the dynamic schedule will perform worse and the logarithmic and uniform schedules will perform better than the standard schedule on the angular resolution (dev) metric (**Hypothesis 7**).

V. Experimental Comparison

This analysis tests the hypotheses developed in the exploratory analysis (Section IV) on a new set of 50 graphs. These graphs are different from the graphs used in Section IV, but they were collected using the same process.

Following existing best practices [30], [31], this evaluation's experimental design is as follows. It uses a within-subjects design on the three schedules (dynamic, logarithmic with $k = 4$, and uniform with $k = 4$), two approximation algorithms (BH and FMM), and four layout readability metrics discussed in Section IV. The first quartile, median, and third quartile for the dynamic schedule's number of updates on the Barnes-Hut algorithm was 11, 12, and 13, and for the Fast Multipole Method algorithm it was 12, 14, and 15. Because this is comparing three alternative schedules to the standard schedule across five response variables (runtime and the four readability metrics), the evaluation uses a Bonferroni corrected significance level $\alpha = 0.05/15 = 0.00\overline{3}$. The corresponding confidence interval is $[0.001\overline{6}, 0.998\overline{3}]$.

Because the readability metrics are bounded in $[0, 1]$, and many are not normally distributed, this evaluation uses bootstrapped confidence intervals with 10,000 samples. Effect sizes

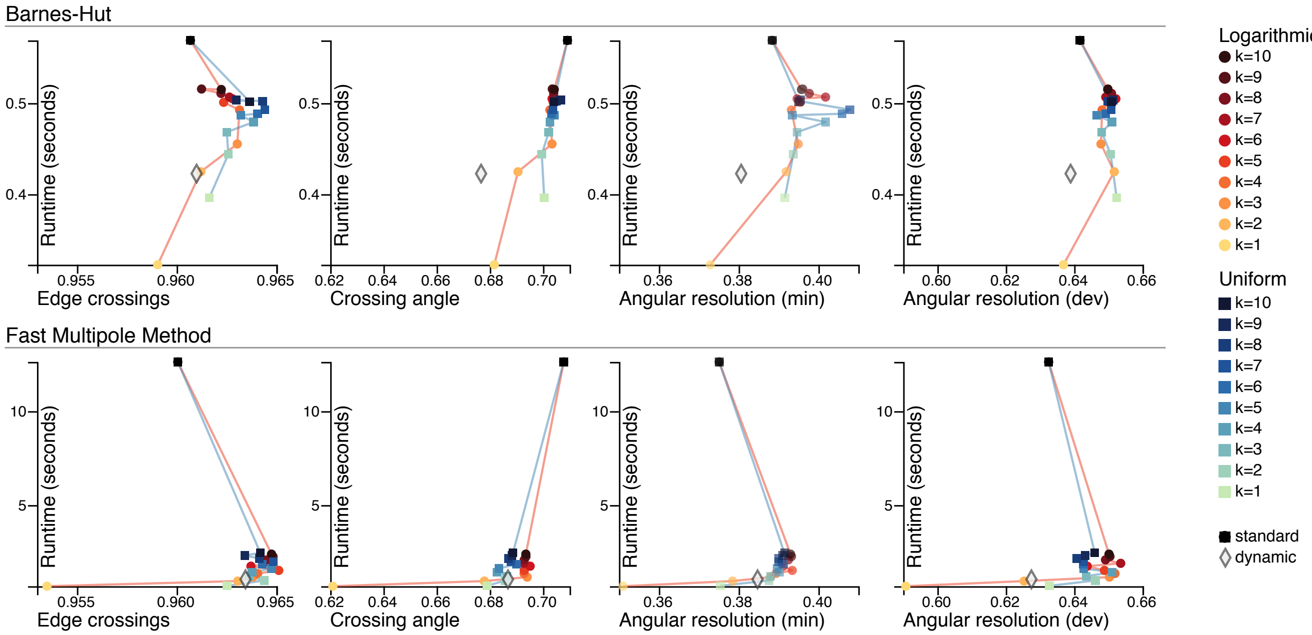


Fig. 1. Median runtime and readability metrics across all 50 exploratory graphs calculated for each update schedule (logarithmic for k from 1 to 10, uniform for k from 1 to 10, the standard schedule, and the dynamic schedule) and approximation algorithm (Barnes-Hut and Fast Multipole Method).

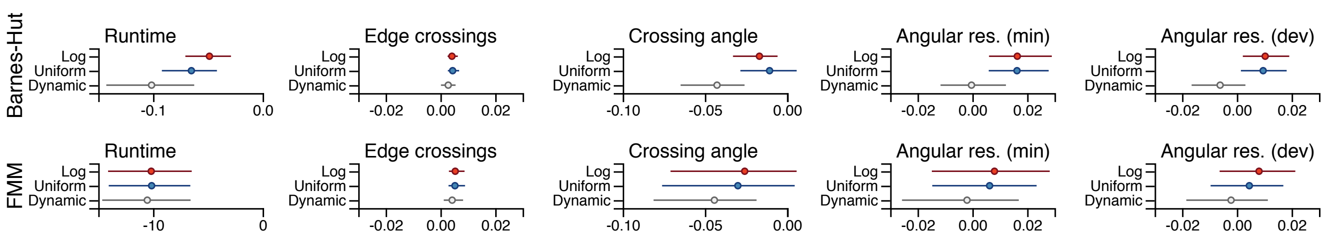


Fig. 2. Bootstrap 99.$\overline{6}$% confidence interval of the effect sizes of the runtime and graph readability metrics. Confidence intervals are percentile bootstrap with 10,000 samples. Effect size is sample mean difference (log, uniform, or dynamic minus standard). Note the differences in x-axis scales.

are sample mean difference ($s_{alt} - s_{standard}$ where s_{alt} is the value for one of the alternate schedules, i.e. logarithmic, uniform, or dynamic). Therefore negative effects indicate the standard schedule has a higher value, whereas positive effects indicate the alternate schedule under test has a higher value. Figure 2 shows the estimation plots for the runtime and readability metrics for the two approximation algorithms.

None of the confidence intervals in the runtime plots cross 0. This indicates that all of the tested update schedules perform faster than the standard update schedule for both the Barnes-Hut (BH) and Fast Multipole Method (FMM) approximation algorithms, which supports **Hypothesis 1**.

Although this experiment does not explicitly test **Hypothesis 2** using a null hypothesis significance test, the dynamic schedule has a much larger effect size on runtime than the logarithmic or uniform schedules for BH. The raw effect size is -0.101 seconds, compared to -0.065 seconds for uniform, and -0.049 seconds for logarithmic. This supports, but does not prove, Hypothesis 2.

We have good evidence to support parts of **Hypothesis 3**. Compared to the standard schedule on BH, the dynamic schedule is better on edge crossings than the standard schedule (the effect is positive and the confidence interval does not cross 0), but the dynamic schedule is worse on crossing angle. For both angular resolution metrics the confidence interval crosses zero, and therefore there is no evidence to say that the dynamic schedule performs better or worse than the standard schedule.

Similarly, we have good evidence to support parts of **Hypothesis 4**. For BH, both the logarithmic and uniform schedules perform better than the standard schedule on edge crossings and the angular resolution metrics. On the crossing angle metric, the logarithmic schedule performs worse than the standard schedule, but the confidence interval for the uniform schedule crosses 0, so we are unable to say whether its performance is better or worse than the standard schedule.

We partly accept **Hypothesis 6**. Although none of the FMM confidence intervals for edge crossings cross 0, they all cross 0 for the angular resolution (min) metric. Therefore we conclude that all of the schedules perform better on edge crossings than the standard schedule, but we cannot say whether there

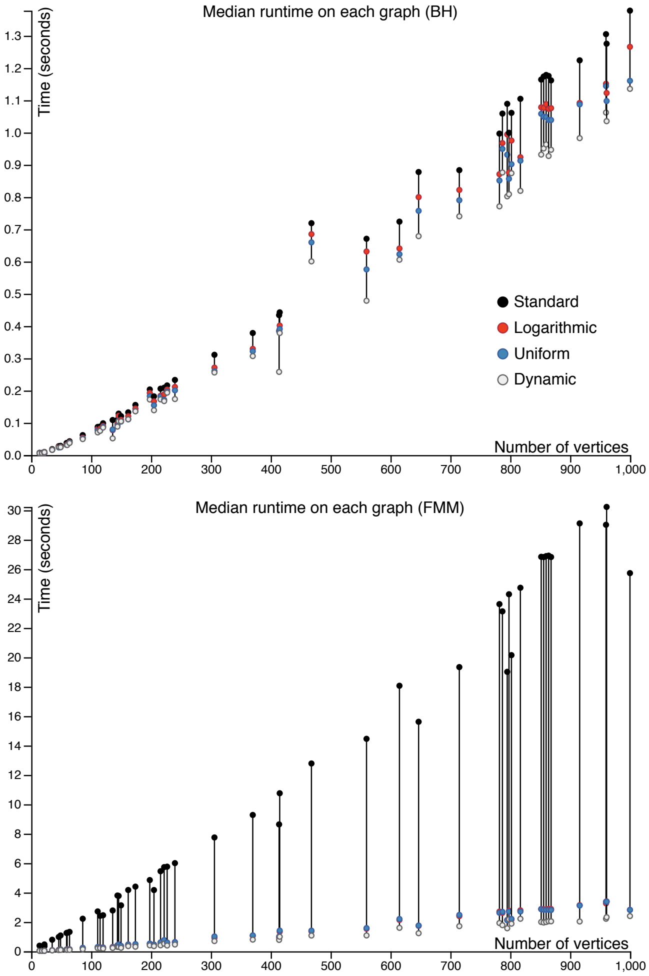


Fig. 3. Median runtime on each of the 50 graphs in the experimental data set. Dots represent the standard schedule (black), the logarithmic schedule (red), the uniform schedule (blue), and the dynamic schedule (gray). The top graph is for the Barnes-Hut algorithm, and the bottom is for Fast Multipole Method. Horizontal jitter resolves occlusion for graphs with identical numbers of vertices.

is any difference from the standard schedule for the angular resolution (min) metric.

Although **Hypothesis 6** is that all of the schedules perform worse than the standard schedule on crossing angle, the evidence only shows that the dynamic schedule performs worse. We do not have enough evidence to say that the logarithmic or uniform schedules perform worse.

Finally, we do not find evidence to support **Hypothesis 7**; all of the confidence intervals for the angular resolution (dev) cross 0, so the logarithmic, uniform, and dynamic schedules do not appear to have significantly different performance from the standard schedule.

VI. Effect of Graph Size on Runtime

The analysis in Section V showed that the logarithmic, uniform, and dynamic schedules have a decrease in runtime over the standard schedule. This section explores this further

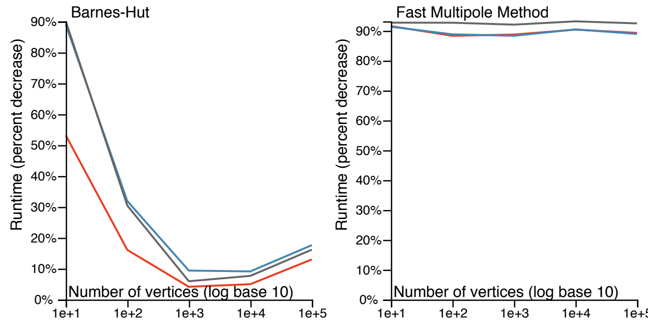


Fig. 4. Percent decrease in runtime from the standard schedule for the logarithmic (red), uniform (blue), and dynamic (gray) update schedules. The percent decrease is calculated for each of the sparse graphs, which have 10, 100, 1,000, 10,000, and 100,000 vertices. Note the log scale on the x-axis.

to better understand the runtime decrease seen in practice and to better understand the influence of graph size.

Figure 3 shows the runtime of each schedule with each approximation algorithm on each of the 50 graphs used in the experimental analysis in Section V.

For Barnes-Hut (BH), the dynamic algorithm has the largest decrease in runtime compared to the standard schedule (the smallest percent decrease is 10%, the largest is 52%, and the median is 18%). (Percent decrease is the difference between the schedule's runtime and the standard schedule's runtime, divided by the standard schedule's runtime.) The uniform schedule did better than the logarithmic schedule in the worst and typical cases (the smallest and median percent decrease was 8% and 14% for the uniform schedule and 5% and 9% for the logarithmic schedule), but in the best case the logarithmic schedule slightly outperformed the uniform schedule (29% compared to 27%).

For the Fast Multipole Method (FMM), the percent decrease in runtime is substantially larger, but very similar for each update schedule (bottom of Figure 3). The logarithmic schedule ranged from 88% to 90% (median 89%), the uniform schedule ranged from 87% to 90% (median 89%), and the dynamic schedule ranged from 89% to 94% (median 92%).

In order to examine the performance of each schedule as $|V|$ increases, Figure 4 shows the percent decrease of each schedule compared to the standard schedule on five sparse graphs of varying size. The five sparse graphs have 10, 100, 1,000, 10,000, and 100,000 vertices, and the graphs were generated with $|E| = 2|V|$ random edges (without replacement). Keeping the proportion of edges fixed allows us to more easily understand the change in runtime performance for repulsive force calculation as graphs get larger.

For BH, the uniform and dynamic schedules have the largest percent decrease for small graphs, but the overall runtime is small so this translates into a savings of about 26 and 37 milliseconds for graphs with 10 and 100 vertices respectively when using the uniform update schedule. For larger graphs, the uniform schedule appears to offer the largest percent decrease; for the largest graph, the percent decrease is about 18%, which saves about 86 seconds (reduced from about 490 seconds to

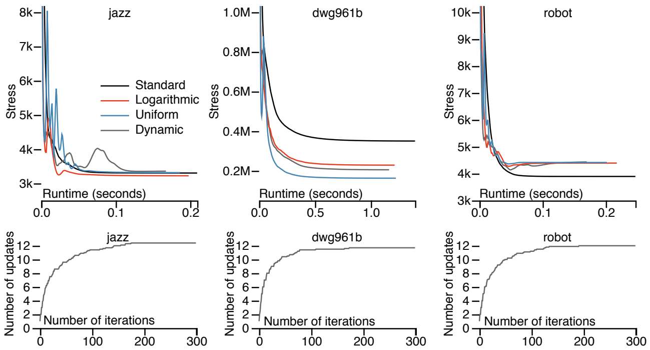


Fig. 5. (TOP) Convergence rate of the stress metric for the four update schedules on the Barnes-Hut approximation. (BOTTOM) The average number of updates of the dynamic schedule at each iteration.

about 404 seconds). For comparison, on the largest graph the dynamic schedule had a percent decrease of about 16%, or about 79 seconds.

For FMM, the percent decrease is large and consistent regardless of graph size and update schedule, ranging from about 88% to about 93%. In contrast to BH, the dynamic schedule has a larger percent decrease, whereas the uniform and logarithmic schedules are about the same.

For BH, the dynamic schedule updated 11 times on the smallest graph and either 8 or 9 times on all other graphs, while for FMM the dynamic schedule updated 18 times on the smallest graph and 12 times on all other graphs. This indicates that the size of the graph does not seem to affect the number of updates.

VII. Convergence

Although using stress as a graph layout quality metric can have problems (see the discussion in Section IV), it is widely used and regarded as an important metric. We are also interested in understanding whether the logarithmic, uniform, and dynamic schedules affect the convergence rate of the layout algorithm. Therefore, this section presents an analysis of stress for each update schedule over 300 iterations on three graphs. These graphs were chosen from the 50 experimental graphs. The stress and layout time are averaged after each iteration across 10 runs of each update schedule. The results are shown in Figure 5. (Due to space constraints, only the Barnes-Hut convergence rates are shown.) Sometimes all of the schedules converge to similar values (the jazz graph), sometimes the alternative schedules converge faster and to lower values than the standard schedule (the dwg961b graph), and sometimes the standard schedule converges to lower values (the robot graph). In most cases, the alternative schedules converge less smoothly than the standard schedule, which is probably because the tree oscillates between being up to date and out of date. However, the logarithmic and uniform schedules have converged by the time the standard schedule has.

The bottom of Figure 5 shows the average number of updates of the dynamic schedule at each iteration. We see that most updates occur during the early iterations. This appears to mirror the convergence charts, and this indicates that the dynamic schedule constructs new trees less often as the layout converges. By the time the layout has converged, the dynamic algorithm typically has stopped constructing new trees.

VIII. Graph Layout Examples

Figure 6 shows the layouts for three graphs generated by the Barnes-Hut (BH) and Fast Multipole Method (FMM) algorithms using the standard, dynamic, logarithmic, and uniform update schedules. These are the same graphs used to test convergence in Section VII. For the most part, the layouts produced by the standard schedule are extremely similar to the layouts produced by the other schedules. Notably, the dynamic FMM layout for the jazz and robot graphs appears to put some vertices too close together while putting other vertices too far apart. The layouts for dwg961b all appear very similar and seem to show the same structure and shape. The dynamic BH and some of the FMM layouts also appear to hide the curvature of the robot layout (due to the central gravitational force in the layout algorithm). These results corroborate the findings in Section V that the logarithmic and uniform schedules produce layouts with similar quality as the standard schedule. It also indicates that the final stress values shown in Figure 5 may not indicate worse layout quality.

IX. Discussion

The analysis in this paper shows that the time required to calculate new trees is a nontrivial part of the overall runtime of spring-electric algorithms that use the Barnes-Hut (BH) approximation and the Fast Multipole Method (FMM). By reducing the number of times the algorithm computes a new tree, we can reduce the time required to compute layouts. Note that these experiments report the reduction in the total running time of graph layout algorithms rather than only the reduction in repulsive force calculation running time. The graph layout algorithms also include calculations for spring forces and a central gravitational force, so the reduction in repulsive (electric) force calculation running time is likely much larger than the results reported here.

Overall, the dynamic schedule appears to be the fastest, although in some cases the runtime improvement may not be much more than the logarithmic and uniform schedules. The dynamic schedule's improvement in runtime appears to come at the cost of lower quality graph layouts. The runtime of the logarithmic and uniform schedules appear similar, although perhaps uniform is slightly faster. Furthermore, the uniform schedule does not appear to have the diminished graph layout quality that we sometimes see in the logarithmic schedule, and in many cases it performs better on readability metrics than the standard schedule. Therefore, practitioners and system implementers would be best choosing either the uniform schedule with $k = 4$ (i.e. construct a new tree approximately once every 13 or 14 iterations) or the dynamic schedule, depending on their preference for trading off speed and quality. These schedules provide modest runtime improvements for BH, but quite substantial runtime improvements for FMM.

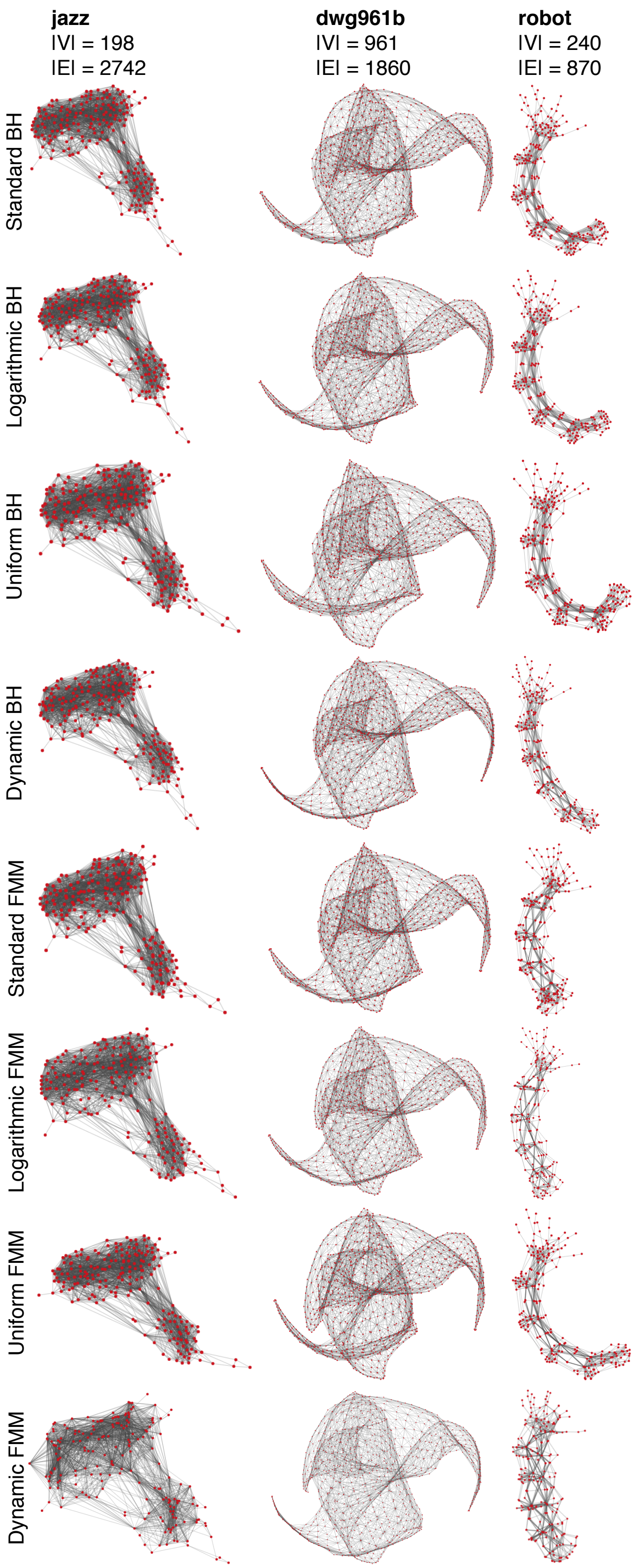


Fig. 6. Layouts of three graphs produced by the Barnes-Hut (BH) and Fast Multipole Method (FMM) algorithms using the four update schedules.

The fact that the uniform schedule is less precise than the standard schedule but produces better quality layouts seems counter intuitive. However, this is consistent with the preliminary results presented by Lipp et al. [5].

The reason why the logarithmic, uniform, and dynamic schedules improve the FMM runtime so much is likely because building the tree for the FMM is $O(|V|\log|V|)$, but computing forces is only $O(|V|)$. Therefore, constructing a new tree is the largest computational cost, and reducing the number of trees constructed has a major impact on the runtime.

It is not clear why the FMM implementation used in this evaluation is so much slower than the BH implementation. The implementations were created by different people, so it is likely that the performance difference is due to differing levels of effort to optimize the implementations. Nonetheless, we still get a clear idea of how much runtime can be reduced by using the alternative update schedules.

It is worth noting that the observed layout quality effect sizes in the experimental analysis may be considered small in practice. For example, an effect size of -0.011 on crossing angle corresponds to a difference of 0.77 degrees in mean deviation from the ideal crossing angle. And an effect size of 0.004 on edge crossings means that if a graph has 1000 edges that can cross, then there will be 4 fewer edge crossings. It is unclear if humans would notice such small differences in practice, or if such small differences would have a detectable effect on speed or errors in analysis tasks.

X. Conclusion

This paper presented two new update schedules (uniform and dynamic) for determining when to construct a new tree in tree-based approximation algorithms such as Barnes-Hut, the Fast Multipole Method, or the Well-Separated Pair Decomposition. The evaluations show that constructing a new tree at a uniform frequency of once every 13–14 iterations achieves significantly faster performance compared to the standard update schedule. In addition, the uniform schedule achieves better edge crossing and angular resolution graph readability metrics, and it does not appear to come at the cost of a degradation in edge crossing angle that occurs with the dynamic and logarithmic update schedules. The uniform update schedule also appears to improve the angular resolution metrics for the Barnes-Hut approximation, but not with the Fast Multipole Method. These new update schedules are simple modifications to the existing approximation algorithms, and therefore they present an easy but effective way to improve the runtime.

Because spring-electric algorithms have many uses, such as in multi-level graph layout algorithms [14] and flow diagrams [32], the logarithmic, uniform, and dynamic schedules can be used to improve runtime in many applications. This includes domains such as n-body simulations or the t-SNE algorithm [33]. Future work should evaluate these other uses.

Acknowledgement

Thanks to Nathan Danneman, Ben Gelman, Jessica Moore, Tony Wong, and the anonymous reviewers for providing helpful feedback on this work.

REFERENCES

[1] A. Arleo, W. Didimo, G. Liotta, and F. Montecchiani, "A Million Edge Drawing for a Fistful of Dollars," in *International Symposium on Graph Drawing and Network Visualization*, 2015, pp. 44–51.

[2] T. M. Fruchterman and E. M. Reingold, "Graph drawing by forcedirected placement," *Software: Practice and Experience*, vol. 21, no. 11, pp. 1129–1164, 1991.

[3] S. G. Kobourov, "Force-Directed Drawing Algorithms," in *Handbook of Graph Drawing and Visualization*, 1st ed., R. Tamassia, Ed. Boca Raton: Chapman and Hall/CRC, 2016, ch. 12, pp. 383–408.

[4] F. Lipp, A. Wolff, and J. Zink, "Faster force-directed graph drawing with the well-separated pair decomposition," in *International Symposium on Graph Drawing and Network Visualization*, 2015, pp. 52–59.

[5] ——, "Faster Force-Directed Graph Drawing with the Well-Separated Pair Decomposition," *Algorithms*, vol. 9, no. 3, p. 53, 2016.

[6] M. Bastian, S. Heymann, and M. Jacomy, "Gephi: An Open Source Software for Exploring and Manipulating Networks," in *International AAAI Conference on Weblogs and Social Media*, 2009, pp. 361–362.

[7] N. G. Belmonte, "JavaScript InfoVis Toolkit." [Online]. Available: http://philogb.github.io/jit/

[8] M. Bostock and J. Heer, "Protovis: A graphical toolkit for visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, pp. 1121–1128, 2009.

[9] M. Bostock, V. Ogievetsky, and J. Heer, "Data-Driven Documents," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2301–2309, 2011.

[10] E. R. Gansner and S. C. North, "An open graph visualization system and its applications to software engineering," *Software Practice and Experience*, vol. 30, no. 11, pp. 1203–1233, 1999.

[11] S. Hachul and M. Jünger, "Large-Graph Layout Algorithms at Work: An Experimental Study," *Journal of Graph Algorithms and Applications JGAA*, vol. 11, no. 2, pp. 345–369, 2007.

[12] J. Heer, S. K. Card, and J. A. Landay, "Prefuse: a toolkit for interactive information visualization," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2005, pp. 421–430.

[13] J. Heer, "Flare: Data Visualization for the Web," 2008.

[14] Y. Hu, "Efficient, High-Quality Force-Directed Graph Drawing," *Mathematica Journal*, vol. 10, no. 1, pp. 37–71, 2005.

[15] M. Jacomy, T. Venturini, S. Heymann, and M. Bastian, "ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software," *PloS one*, vol. 9, no. 6, 2014.

[16] M. Khoury, Y. Hu, S. Krishnan, and C. Scheidegger, "Drawing Large Graphs by Low-Rank Stress Majorization," *Computer Graphics Forum*, vol. 31, no. 3pt1, pp. 975–984, 2012.

[17] J. Nielsen, *Usability Engineering*, 1st ed. Academic Press, 1993.

[18] J. Barnes and P. Hut, "A hierarchical O(N logN) force calculation algorithm," *Nature*, vol. 324, pp. 446–449, 1986.

[19] A. Quigley and P. Eades, "FADE: Graph drawing, clustering, and visual abstraction," in *International Symposium on Graph Drawing*, 2000, pp. 197–210.

[20] S. Aluru, J. Gustafson, G. Prabhu, and F. E. Sevilgen, "Distribution-independent hierarchical algorithms for the n-body problem," *The Journal of Supercomuting*, vol. 12, no. 4, pp. 303–323, 1998.

[21] L. F. Greengard, "The rapid evaluation of potential fields in particle systems," Ph.D. dissertation, Yale University, New Haven, CT, USA, 1987.

[22] S. Hachul and M. Jünger, "Drawing large graphs with a potential-field-based multilevel algorithm," in *International Symposium on Graph Drawing*, 2004, pp. 285–295.

[23] P. B. Callahan and S. R. Kosaraju, "A decomposition of multidimensional point sets with applications to k-nearest-neighbors and n-body potential fields," *Journal of the ACM*, vol. 42, no. 1, pp. 67–90, 1995.

[24] J. Kunegis, "KONECT - The Koblenz Network Collection," in *Web Observatory Workshop*, 2013, pp. 1343–1350.

[25] T. Davis and Y. Hu, "The University of Florida Sparse Matrix Collection," *ACM Transactions on Mathematical Software*, vol. 38, no. 1, pp. 1:1–1:25, 2011.

[26] P. Eades, S.-H. Hong, A. Nguyen, and K. Klein, "Shape-Based Quality Metrics for Large Graph Visualization," *Journal of Graph Algorithms and Applications*, vol. 21, no. 1, pp. 29–53, 2017. [Online]. Available: http://jgaa.info/getPaper?id=405

[27] J. F. Kruiger, P. E. Rauber, R. M. Martins, A. Kerren, S. Kobourov, and A. C. Telea, "Graph Layouts by t-SNE," *Computer Graphics Forum*, vol. 36, no. 3, pp. 283–294, 2017.

[28] E. R. Gansner, Y. Koren, and S. C. North, "Graph drawing by stress majorization," in *International Symposium on Graph Drawing*, 2004, pp. 239–250.

[29] C. Dunne, S. I. Ross, B. Shneiderman, and M. Martino, "Readability metric feedback for aiding node-link visualization designers," *IBM Journal of Research and Development*, vol. 59, no. 2/3, pp. 14:1–14:16, 2015.

[30] P. Dragicevic, "Fair statistical communication in hci," in *Modern Statistical Methods for HCI*, J. Robertson and M. Kaptein, Eds. Springer, 2016, ch. 13, pp. 291–330.

[31] C. Rainey, "Faster Force-Directed Graph Drawing with the Well-Separated Pair Decomposition," *Journal of Political Science*, vol. 58, no. 4, pp. 1083–1091, 2014.

[32] K. Wongsuphasawat and D. Gotz, "Exploring flow, factors, and outcomes of temporal event sequences with the outflow visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2659–2668, 2012.

[33] L. van der Maaten, "Accelerating t-sne using tree-based algorithms," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 3221–3245, 2014.

Session 3: HCI

ViReSt - Storytelling with Volumetric Videos

Gerhard Sprung
Dep. Information Management
FH JOANNEUM
Graz, Austria
Gerhard.sprung@fh-joanneum.at

Alexander Nischelwitzer
Dep. Information Management
FH JOANNEUM
Graz, Austria
Alexander.nischelwitzer@fh-joanneum.at

Andreas Egger
Dep. Information Management
FH JOANNEUM
Graz, Austria
Andreas.egger@edu.fh-joanneum.at

Robert Strohmaier
Dep. Information Management
FH JOANNEUM
Graz, Austria
Robert.strohmaier@fh-joanneum.at

Sandra Schadenbauer
Dep. Information Management
FH JOANNEUM
Graz, Austria
Sandra.schadenbauer@fh-joanneum.at

Abstract
The aim of the project ViReSt is to explore the extent to which virtual reality, volumetric video and binaural audio can be utilized to create an immersive experience for museum visitors. Museum curators should be able to develop, test and create the story themselves without programming skills. Therefore, a workflow was developed and tested by creating a small prototype. This prototype was then tested with a group of 10 participants to assess immersion, reception of volumetric video and binaural audio.

Keywords — VR; volumetric video; storytelling

I. INTRODUCTION

In the near future Virtual Reality (VR) will be mature enough to be widely used in museums. The project ViReSt (VIrtual REality StoryTelling) addresses the question of how to create compelling content for such projects. With our prototype, we explored what tools are or would be needed to give museum designers without programming knowledge the ability to create interactive stories that are immersive, emotionally engage visitors and provide a level of interactivity.

The “wow-effect”, which is usually sufficient to attract the visitors’ attention to an exhibit that is presented with VR, will no longer suffice as soon as the visitors will already have experience with VR-devices. In the same way as in the advent of film and again in the early stages of the computer game industry, traditional methods and approaches for creating immersive experiences have to be rethought, reworked, redefined or even created from scratch [1].

II. STATE OF ART

A. Storytelling in Museums

To offer immersive experiences in museums, several technologies are currently to be considered. Audio guides with radio plays and spoken stories are already widely used [2][3] although interactivity is mostly limited to starting and stopping the recording. There are also approaches where the contents are location-dependent [5] or the visitors can select a role from whose perspective the story is told [6]. Films are often used, sometimes presented with methods that increase immersion like the Pepper’s Ghost Effect[1], projection mapping, projections on spheres and dome projection.

Museums try to personalize the visitor’s experience, make it interactive and adapt it to interests and needs. The personalization of the experience during a museum visit is possible but brings along various technical and content-related problems: in many museums, it not feasible for the exhibit to be viewed by only one person at a time [3]. This raises the problem that everyone sees and hears the same thing. Even if the acoustic information can be perceived independently with the help of headphones or sound domes [7], the obstacle still arises that the playing system must be aware of the identity of the visitor, including their previously chosen role [6], the information

[1] The Pepper’s Ghost effect is an optical illusion that utilizes large pieces of glass to partly reflect a scene which cannot be directly observed by the audience. The viewers see the reflections as “ghosts”.

already received and the path through the exhibition so far. This data has then to be used to calculate and coordinate the trajectory and distinct information that is presented [8]. However, if the information should also be optically adapted to the role and trajectory of every particular visitor, this can only be done by spatial separation like using separate rooms, booths or head mounted displays.

B. Characters in VR

If a museum uses storytelling to touch visitors emotionally, they have to provide them with believable and convincing protagonists to interact with.

Several researchers found that high behavioral realism and realistic visualization of the characters in a movie increase the believability [9]–[12]. But on the other hand, artists [13] and other researcher stress the problem of the "Uncanny Valley-Effect" [14]. This effect describes the phenomenon that the audience find small discrepancies in the behavior or the realistic visualization of extremely realistic characters very disturbing (often described as "eerie") [14][15][16] whereas a cartoony stylization is very forgiving and can concentrate on the personality of a character [17][18].

Concerning the creation and digital representation of the virtual actors there are several technical possibilities.

1) Mesh and Bones

The surface of a digital character is stored digitally and, with the help of shaders (instructions how to calculate the color of each pixel) and light simulations, an image is rendered. Game engines are capable of creating very realistic scenes in real-time.

To create such so called "meshes" artists can use Computer Graphics (CG) programs to construct and sculpt the surface. Alternatively, real-life persons can be measured and digitized. These values can be used to reconstruct the person as a mesh. One wide-spread approach to achieve this is by scanning with a laser, infrared (IR) or structured light. Another approach is the calculation of the mesh with the help of Photographs and Structure from Motion (SfM) or other photogrammetry techniques.

To animate these meshes usually digital bones are constructed to provide an underlying structure which is utilized to deform the skin. The movement of these bones either can be animated by an animator (this is known as "keyframe animation") or is acquired from human motion with the help of motion capture techniques.

2) Images and Movies (Billboards)

A second approach is to use 360° movies, often called spherical or 360/180 movies. In this approach, real-life scenes are filmed with a special camera to record a panoramic view. Therefore, creating an interactive application with this technique requires recording all possible versions. If the user makes a decision, the movie has to switch to the corresponding variation.

It is possible to combine spherical movies with 3D-representation. Here the spherical movie is projected onto the background and the characters are placed in 3-D around the virtual camera. Characters can be represented as 3-D-meshes as described above or as 2-dimensional objects showing a movie (this is called a "billboard"). In this case, the shown characters can be life-action footage or stylized representations.

Such films can then be viewed with Head Mounted Displays (HMDs) but, apart from the head rotation, they do not allow much interaction and are accordingly rarely immersive experiences.

3) Volumetric Video

In order to facilitate Free Viewpoint Video (FVV), techniques are developed to store video in a way that the viewer can choose his viewpoint arbitrarily. Volumetric Video allows the recording of people and their movement utilizing the depth information of depth-cameras ("RGBD-camera") [19][20], photogrammetry or lasers. To store the results, the surface is either represented as a point cloud[2] [21][22] for each frame of an animation, a dynamic mesh including all movement of the vertices[3] or the depth information can be stored separately as an image. In this way it is possible to store animations without using bones and skin deformations [23][24].

C. Story Planning for VR

1) Planning

In the film industry, the description of a film in form of a movie script is highly standardized. To prevent ambiguities and enable effective cooperation scripts are expected to have a special form and structure. Another tool for planning a film, Storyboards, are visualizations of each shot of the movie. Storyboards and the predecessor of the modern movie script (Continuity Scripts) have been used since 1928 [25][26]. With the help of storyboards, one can virtually watch the movie by looking on the images in the right order.

Neither of these methods is applicable for 360° video, VR or interactive storytelling. Since the viewer can choose their perspective and their position in the scene as well as influence the story new techniques to plan the experience are needed [27] [28].

Because of the many different ways a non-linear story can unfold, a system to previsualize an interactive story has to be capable of using variables, conditions and user input. Also it is necessary to deal with the distance to the viewer, their possible positions and what the viewer sees [28].

Interactive story editors like Twine[4] are designed to create text-based stories. Tools like articy:draft[5] (Figure 1) are already used to create and organize interactive and non-linear game content. Programs that allow drawing, modeling and animation directly in VR [29][30][31], as well as manipulation of the flow of an interactive application [32] help to position and control

[2] A point cloud is a set of points in space

[3] A vertex is a point in space used to define a mesh

[4] http://twinery.org/2

[5] https://www.nevigo.com/en/articydraft/

visual cues. This prevents situations where the viewers oversee important clues for the understanding of the story.

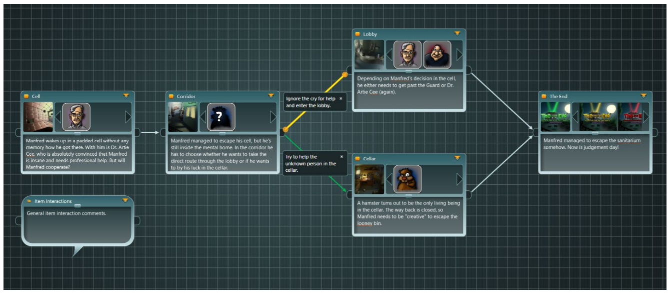


Figure 1: Articy:draft

Although there are tools to support the development of interactive content, real interactive utilities are still missing in order to clearly present and manipulate the complex processes as well as provide visual aids and automatically check for weak points and dead ends.

2) Directing

In traditional movies the director guides the attention of the audience by choosing an appropriate frame and therefore restricts the perceptional boundary. The director can decide upon the position of the camera (Point of View), the focus, movement of camera and actors and so forth. In 360°-movies and VR-applications the viewer himself to a high extent is choosing the point of view. Although there is important research being done concerning the use of focus and depth of field in VR, these possibilities are not yet widely available [33].

In order to increase immersion and to understand the story it is extremely important to make viewers look and move in the right direction at specific times. Therefor the director has to use visual [34] and acoustic [27] cues to help the audience to understand what is happening and to guide them through the story.

If the viewers are able to move freely in the scene, their possibilities should be restricted by obstacles (like walls or fences) in the virtual world. In this case, the visitors have to be provided with methods to cover larger distances. Flying and being moved in the virtual world often leads to nausea because of the discrepancies between real movement and the movement shown in the VR-Application[35][36]. This effect seems to be independent of the amount of immersion [37]. One promising approach to tackle this problem is teleporting. By pointing at the point where one wants to be the viewer instantly changes position [38][39][40]. Sometimes stories need different ways of navigating like a wheelchair[41].

One important help to orient in a virtual world is binaural audio [27]. By manipulating the frequencies of an audio event, the frequency change occurring in the outer ear (Head Related Transfer Function HRTF) can be simulated [42][43]. Therefor the sound seems to be coming not only from the side but also from above, below or behind [44]. The sound sources are placed at the appropriate positions in the virtual space. The sound is manipulated depending on the position of the viewer, the direction of the head and the size and surface of the room and obstacles, and send to the headphones. Thus the user can recognize if a sound event is occurring e.g. behind or above him [45].

With the help of these technologies it is possible to create the basic conditions for the so-called immersion or presence [46].

III. Proof-of-Concept

Based on this research we developed a workflow to create interactive stories incorporating volumetric video, spatial sound, interaction triggered by various sensors and events and speech input.

A. Planning

For the first phase –planning- we experimented with different tools like Twine, several script editing Tools like celtX[6], scrivener[7] and Trelby[8]. Additionally storyboarding tools like storyboarder[9] and articity:draft were evaluated. The highly branched structure[47][48], which is also dependent on conditions and variables has led us to exclude all script-editors and linear storyboarding tools. In future projects we will use Twine, a simple Editor for interactive text stories which is capable of visualizing the structure and the flow of the story (Figure 2). With this software, even variables and conditions can be defined and interactively tested.

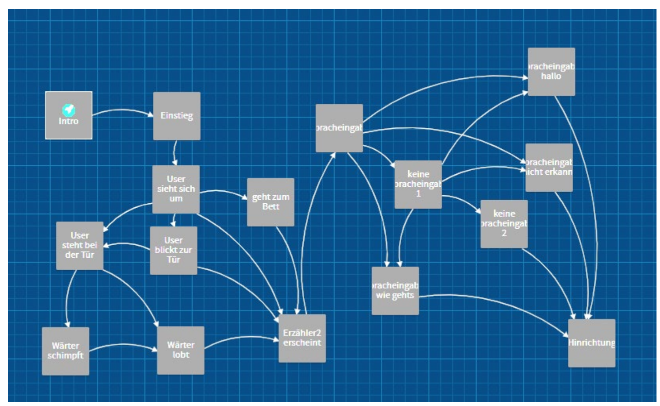


Figure 2: Story structure in Twine

B. Recording Volumetric Video

The method we have chosen to employ for our proof of concept is volumetric video. To allow a simple volumetric representation, we restricted the representation to a 2.5-D representation that spatially visualizes about half of the human body. Our recording software encodes both the depth and color information of a RGBD-Camera as two RGB-images into a movie file. One image containing the color information and one image the depth information are stored side-by-side (Figure 3).

[6] https://www.celtx.com/index.html

[7] https://www.literatureandlatte.com/scrivener/overview

[8] https://www.trelby.org/

[9] https://wonderunit.com/storyboarder/

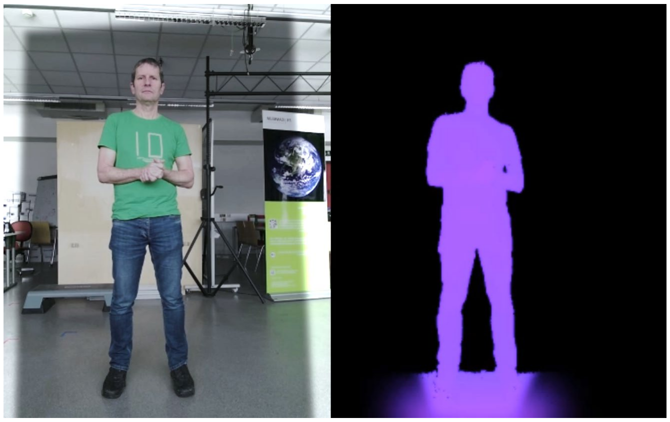


Figure 3: Side by side film

We used the Microsoft Kinect 2, which has 512x424-pixel resolution for the depth information, which was sufficient for our tests. To improve the results, we processed the images with the post-production software Blackmagic Fusion[10]. We extracted the background, improved color and contrast, softened the edges and enhanced the depth information.

C. Editor

To add interactivity, we had to develop an editing application that was capable of defining objects, events, sensors and triggers (Figure 4).

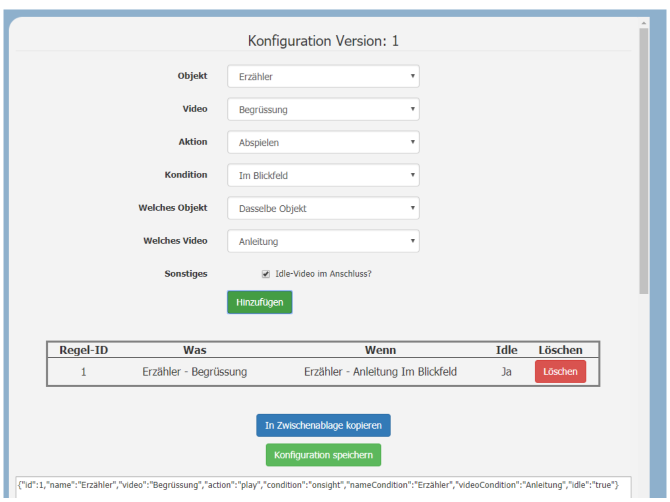


Figure 4: Screenshot of the Editor

Objects can play animations and sound files if they are triggered and all conditions are met. Therefore, features for defining variables, timers and possibilities for user-input had to be provided [49]. With this editor, even persons who are not used to writing code are able to design complex interactive scenes and branched plots. The resulting definitions of scenes can be tested in a generic 3D-environment to detect any gaps, loops or dead ends of the story and to test different paths through the story.

For future applications we will integrate visual editing. For this purpose, screenshots will be created and displayed in the 3D environment. The connections between events will be displayed as lines and the corresponding conditions and variables will be visualized.

We also intend to introduce interactive story design in a VR-environment.

D. Player

Within the game-engine Unity3D[11], a mesh is automatically prepared utilizing the previously described video containing the depth information (Figure 5). This is achieved by the help of our vertex-shaders, which refine ("tessellate") the mesh and displace the vertices according to the brightness of the depth-image stored in the movie. Thus a 3-dimensional object is created and shown in the scene.

These volumetric videos are placed in the appropriate positions according to a file created by the editor. The runtime-module (player) displays, starts and ends the presentation of these videos depending on the variables in the scene description. It also reacts to user inputs to trigger processes and animations. The triggers that are implemented in the application are capable of responding to gaze (is the viewer actually seeing an object), time events (has something already happened), and the end of other animations. One additional experimental trigger was the human voice. The Windows Speech API which is available in Unity3D provides the required phrase recognition capability[12]. We implemented the possibility that under certain conditions the user can talk to the character in the scene. The engine searches for predefined phrases in the spoken text and reacts accordingly, e.g. by starting an animation.

To implement binaural audio in our application we used the Resonance Audio SDK by Google[13]. Resonance Audio allows to apply head-related transfer functions (HRTFs) to the soundfield to enable spatial hearing.

As the space that can be detected by the trackers of the Oculus Rift HMD is limited, the space of the story had to be limited accordingly. This was achieved by creating a scene in a prison cell. The walls of the cell represent the borders of the space, which could be tracked. It would have been possible to put real walls where the users saw walls just as it would have been possible to put real chairs and a bed into the scene. The downside of this approach is that it would require additional effort for calibrating the trackers.

[10] https://www.blackmagicdesign.com/at/products/fusion/

[11] https://unity3d.com/

[12] https://developer.microsoft.com/en-us/windows/speech

[13] https://developers.google.com/resonance-audio/

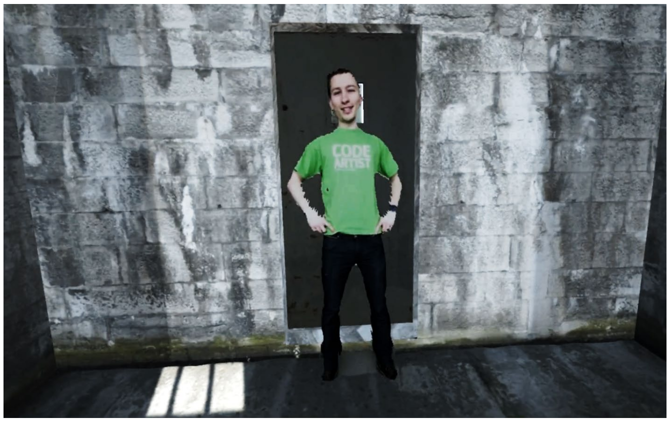


Figure 5: Screenshot of the application

IV. TEST

To assess immersion, reception of volumetric video and of binaural audio we conducted user tests on 10 participants at the ages of 20-26. In a prequestionnaire, the former experiences of the participants concerning VR and interactive applications in museums were surveyed.

In the user test, the participants had to explore an interactive VR-scene. With a HMD (HTC Vive) and headphones including a microphone they were able to experience a scene in a prison cell as a prisoner and interact with other persons in the scene. All actions and reactions of the participants were recorded. With a postquestionnaire, they were questioned about their understanding of the situation and their role in the virtual scene, as well as their emotional involvement and satisfaction.

In the second part of the test, we wanted to test the reception of and reactions to different ways of presenting the volumetric videos. The participants had to rate different versions of visualizations of human characters in the VR-application (Figure 6).

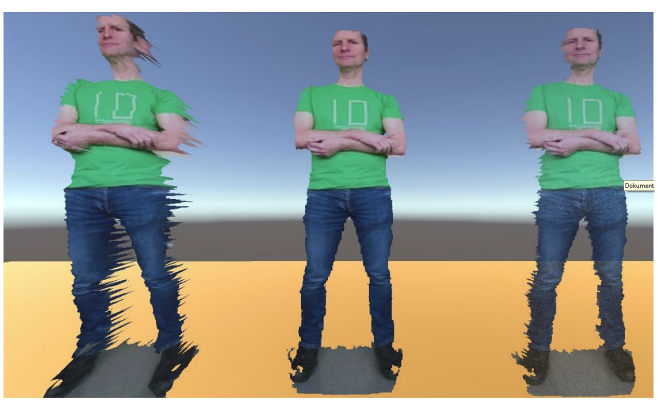


Figure 6: different versions of visualization

V. RESULTS

Ninety percent of the participants were able to identify their role as a prisoner without problems, one participant thought to be an uninvolved observer in the scene. All participants agreed that they were emotionally more involved as they would be in a traditional film. They were able to empathize with the prisoner and understand his feelings. The use of volumetric video was received very positively.

All participants recognized the binaural sound and were positively surprised. They reacted to the acoustic clues and turned exactly into the right direction, often moving toward the perceived source to explore it.

The results of the second part of the test showed that 80% of the participants favored a clear flat representation over more three-dimensional representation. 60% preferred a version where the virtual person automatically faces the viewer ("billboard"). In the last test, the participants were confronted with differently sized characters. All participants reported feeling as intimidated or frightened by the oversized characters. It has to be mentioned that the height of the camera in the scene is dependent on the body height of the user because the sensors of the HMD gauge its exact position and transfer it to the virtual Camera.

One problem, which was not unexpected, was the participants' hesitation to speak loudly and to be seen publicly when wearing a HMD. These problems have to be taken into account when creating such applications for museums. One solution could be to provide small rooms or booths where the visitors can explore the application in private.

We were able to show one possible workflow for creating interactive VR-applications without programming knowledge. In spite of many technical problems that have to be solved to achieve compelling results we were able to realize a small project using volumetric video and an easy to use editor. The first user tests suggest that the approach seems to be very promising concerning the emotional involvement and therefore supporting storytelling in museums.

VI. BIBLIOGRAPHY

[1] S. Kirchhof, "Storytelling for Volumetric VR," in *Proceeding SIGGRAPH '18 ACM SIGGRAPH 2018 Educator's Forum.*

[2] L. Fosh, S. Benford, S. Reeves, B. Koleva, and P. Brundell, "See Me, Feel Me, Touch Me, Hear Me," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. - CHI '13*, p. 149, 2013.

[3] L. Pujol, M. Roussou, S. Poulou, O. Balet, M. Vayanou, and Y. Ioannidis, "Personalizing interactive digital storytelling in archaeological museums: the CHESS project," *Archaeol. Digit. Era. Pap. from 40th Annu. Conf. Comput. Appl. Quant. Methods Archaeol.*, 2013, unpublished.

[4] M. Vayanou, A. Katifori, M. Karvounis, V. Kourtis, M. Kyriakidi, M. Roussou, M. Tsangaris, Y. Ioannidis, O. Balet, T. Prados, J. Keil, T. Engelke, and L. Pujol, "Authoring Personalized Interactive Museum Stories," pp. 37–48, 2014.

[5] J. Alvermann, "Mobile media in the museum space," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct - UbiComp '16*, 2016, pp. 1509–1512.

[6] S. Benford, G. Giannachi, B. Koleva, and T. Rodden, "From Interaction to Trajectories: Designing Coherent Journeys Through User Experiences," *Chi2009 Proc. 27th Annu. Chi Conf. Hum. Factors Comput. Syst. Vols 1-4*, pp. 709–718, 2009.

[7] J. Kuutti, J. Leiwo, and R. Sepponen, "Local Control of Audio Environment: A Review of Methods and Applications," *Technologies*, vol. 2, no. 1, pp. 31–53, 2014.

[8] K. Salo, V. Zinin, M. Bauters, and T. Mikkonen, "Modular Audio Story Platform for Museums," *Proc. 22Nd Int. Conf. Intell. User Interfaces Companion*, pp. 113–116, 2017.

[9] W. Hai, N. Jain, A. Wydra, N. M. Thalmann, and D. Thalmann, "Increasing the feeling of social presence by incorporating realistic

interactions in multi-party VR," *Proc. 31st Int. Conf. Comput. Animat. Soc. Agents - CASA 2018*, pp. 7–10, 2018.

[10] E. Kokkinara and R. McDonnell, "Animation realism affects perceived character appeal of a self-virtual face," *Proc. 8th ACM SIGGRAPH Conf. Motion Games - SA '15*, pp. 221–226, 2015.

[11] M. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch, "The Effect of Avatar Realism in Immersive Social Virtual Realities," in *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, 2017.

[12] L. Freina, R. Bottino, and M. Tavella, "From e-learning to VR-learning: An example of learning in an immersive virtual world," *J. E-Learning Knowl. Soc.*, vol. 12, no. 2, pp. 101–113, 2016.

[13] R. McDonnell, "Appealing virtual humans," *Motion in Games*, vol. 7660, pp. 102–111, 2012.

[14] M. Mori, "The Uncanny Valley," 1970. [Online]. Available: https://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley. [Accessed: 08-Nov-2018].

[15] R. Wages, S. M. Grünvogel, and B. Grützmacher, "How Realistic is Realism? Considerations on the Aesthetics of Computer Games," vol. 3166, pp. 216–225, 2004.

[16] T. Chaminade, J. Hodgins, and M. Kawato, "Anthropomorphism influences perception of computer-animated characters' actions," *Soc. Cogn. Affect. Neurosci.*, vol. 2, no. 3, pp. 206–216, 2007.

[17] F. Thomas and O. Johnston, *The Illusion of Life: Disney Animation*. New York, NY, USA: Hyperion, 1995.

[18] K. F. MacDorman, R. D. Green, C. C. Ho, and C. T. Koch, "Too real for comfort? Uncanny responses to computer generated faces," *Comput. Human Behav.*, vol. 25, no. 3, pp. 695–710, 2009.

[19] "Microsoft Kinect." [Online]. Available: https://developer.microsoft.com/en-us/windows/kinect. [Accessed: 19-Sep-2018].

[20] "Brekel Pro Pointcloud." [Online]. Available: https://brekel.com/. [Accessed: 19-Sep-2018].

[21] Z. N. Sultani and R. F. Ghani, "Kinect 3D Point Cloud Live Video Streaming," *Procedia Comput. Sci.*, vol. 65, pp. 125–132, Jan. 2015.

[22] M. Chuang, P. Sweeney, D. Gillett, D. Evseev, H. Hoppe, A. Kirk, S. Sullivan, and D. Calabrese, "High-Quality Streamable Free-Viewpoint Video," *ACM Trans. Graph.*, vol. 34, no. 4, p. 69, 2015.

[23] "8i." [Online]. Available: http://8i.com/. [Accessed: 19-Sep-2018].

[24] "Microsoft Mixed reality capture." [Online]. Available: https://www.microsoft.com/en-us/mixed-reality/capture-studios. [Accessed: 19-Aug-2018].

[25] C. Finch, *The Art of Walt Disney*, 2nd ed. New York, NY, USA: Portland House, 1988.

[26] J. Hart, *The art of the storyboard: A filmmaker's introduction*. CRC Press, 2013.

[27] P. Jayesh, S, I. Azif, and C. Herold, P, "Grammar of VR Storytelling : Visual Cues," in *VRIC '17 Proceedings of the Virtual Reality International Conference - Laval Virtual 2017*, 2017, pp. 1–4.

[28] M. Alger, "Visual Design Methods for Virtual Reality," 2015. [Online]. Available: http://aperturesciencellc.com/vr/VisualDesignMethodsforVR_MikeAlger.pdf. [Accessed: 08-Nov-2018].

[29] "gravitysketch." [Online]. Available: https://www.gravitysketch.com/. [Accessed: 20-Sep-2018].

[30] "AnimVR." [Online]. Available: https://nvrmind.io/. [Accessed: 20-Sep-2018].

[31] "quill." [Online]. Available: https://quill.fb.com/. [Accessed: 20-Sep-2018].

[32] "spatial stories." [Online]. Available: http://spatialstories.net/). [Accessed: 20-Sep-2018].

[33] L. Shi, F.-C. Huang, W. Lopes, W. Matusik, and D. Luebke, "Near-eye light field holographic rendering with spherical waves for wide field of view interactive 3D computer graphics," *ACM Trans. Graph.*, vol. 36, no. 6, pp. 1–17, 2017.

[34] L. T. T. Nielsen, M. B. B. Moller, S. D. D. Hartmeyer, T. C. M. C. M. Ljung, N. C. C. Nilsson, R. Nordahl, and S. Serafin, "Missing The Point: An Exploration of How to Guide Users' Attention During Cinematic Virtual Reality," *Proc. 22nd ACM Conf. Virtual Real. Softw. Technol.*, pp. 229–232, 2016.

[35] N. Norouzi, G. Bruder, and G. Welch, "Assessing vignetting as a means to reduce VR sickness during amplified head rotations," *Proc. 15th ACM Symp. Appl. Percept. - SAP '18*, pp. 1–8, 2018.

[36] S. Davis, K. Nesbitt, and E. Nalivaiko, "A Systematic Review of Cybersickness," in *IE2014 Proceedings of the 2014 Conference on Interactive Entertainment*, 2014.

[37] S. von Mammen, A. Knote, and S. Edenhofer, "Cyber sick but still having fun," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology - VRST '16*, 2016, pp. 325–326.

[38] E. Bozgeyikli, A. Raij, S. Katkoori, and R. Dubey, "Point & Teleport Locomotion Technique for Virtual Reality," *Proc. 2016 Annu. Symp. Comput. Interact. Play - CHI Play '16*, pp. 205–216, 2016.

[39] J. Liu, H. Parekh, M. Al-Zayer, and E. Folmer, "Increasing Walking in VR using Redirected Teleportation," *Uist 2018*, pp. 521–529, 2018.

[40] N. Coomer, S. Bullard, W. Clinton, and B. Williams-Sanders, "Evaluating the effects of four VR locomotion methods," *Proc. 15th ACM Symp. Appl. Percept. - SAP '18*, pp. 1–8, 2018.

[41] J. Cho, T.-H. Lee, J. Ogden, A. Stewart, T.-Y. Tsai, J. Chen, and R. Vituccio, "Imago," *ACM SIGGRAPH 2016 VR Village - SIGGRAPH '16*, pp. 1–2, 2016.

[42] Y. Suzuki, D. Brungart, Y. Iwaya, K. Iida, D. Cabrera, and H. Kato, *Principles and applications of spatial hearing*. Singapore: World Scientific Publishing, 2011.

[43] J. Villegas and M. Cohen, "HRIR~: modulating range in headphone-reproduced spatial audio," in *Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry*, 2010, vol. 1, no. 212, pp. 89–94.

[44] W. Zhang, P. Samarasinghe, H. Chen, and T. Abhayapala, "Surround by Sound: A Review of Spatial Audio Recording and Reproduction," *Appl. Sci.*, vol. 7, no. 6, p. 532, 2017.

[45] C. H. Larsen, D. S. Lauritsen, J. J. Larsen, M. Pilgaard, and J. B. Madsen, "Differences in human audio localization performance between a HRTF- and a non-HRTF audio system," *Proc. 8th Audio Most. Conf. - AM '13*, pp. 1–8, 2013.

[46] B. G. Witmer and M. J. Singer, "Measuring Presence in Virtual Environments: A Presence Questionnaire," *Presence Teleoperators Virtual Environ.*, vol. 7, no. 3, pp. 225–240, Jun. 1998.

[47] C. Crawford, *On interactive storytelling*, 2nd ed. New Riders, 2012.

[48] M.-L. Ryan, *Narrative as virtual reality: Immersion and interactivity in literature and electronic media*. Baltimore: The Johns Hopkins University Press, 2003.

[49] N. Szilas, "Structural Models for Interactive Drama," *Proc. 2nd Conf. Comput. Semiot. Games New Media (COSIGN 2002)*, pp. 22–27, 2002.

Visual forms of presentation of investigative online journalism in Austrian media

Stefanie Braunisch
Editing, Topics addendum
Vienna, Austria
s.braunisch@hotmail.com

Michael Roither
Information Technology and Management
University of Applied Sciences Burgenland
Eisenstadt, Austria
michael.roither@fh-burgenland.at

Michael Zeiller
Information Technology and Management
University of Applied Sciences Burgenland
Eisenstadt, Austria
michael.zeiller@fh-burgenland.at

***Abstract*- This article deals with the current visual and structural forms of presentation of Austrian media for investigative content. On a theoretical basis the thematic characteristics and working methods of investigative journalism are explained, as well as the basis of online journalism. The most important and most common visual and structural forms of presentation of current online formats are considered for content analysis. The study examines how investigative findings are processed for online presentation using the example of historically or self-defined investigative media. The results of a study among Austrian journalists are presented which identify how those journalists actually use different journalistic text styles, media elements and visual representations to tell their investigative stories. Text is the dominant form of presentation (86 %) used in investigative reports, but pictures (64 %) and information graphics (50 %) are frequently used as well. Video and audio are seldom used to present investigative content. Similar to the actual usage of the analyzed forms of presentation in published reports all journalists consider text to be most important (100 %), followed by information graphics (78 %) and pictures (76 %).**

Keywords—investigative journalism; online journalism; visual journalism; journalistic forms of presentation; text styles; visual representation

I. INTRODUCTION

The Internet has fundamentally changed the communication processes of the media world and makes it possible to use far more diverse forms of presentation and types of use for the benefit of readers than classic print products, for example. As a result, the media should also address this change and create web presences that offer added value in terms of both content and appearance. Since investigative journalism is a particularly research- and fact-intensive form of journalism, content is complex, difficult to understand and as a result often visually more complex to prepare [1]. Due to its social relevance, adequate communication of the content of investigative journalism would be particularly important, as this is the only way to guarantee readers' understanding of highly relevant topics. Online presences with multiple multimedia and interactive elements make it easier to implement this complex content in an understandable and comprehensible manner.

At the same time, the triumphant advance of the Internet has changed the revenue models of many media. Advertising revenues have often fallen drastically as a result, but readers can rarely be persuaded to pay for journalism. One of the more frequently proposed solution concepts is therefore the focus on quality journalism, in which investigative journalism in particular can offer added value [2]. Nevertheless, online journalism is still only used by many media as a second division for the reuse of content, as the special formats of presentation require more resources: more lead time and often also more investments in implementation tools and personnel [3].

This paper investigates how Austrian media that publish investigative content use visual and structural forms of presentation (i.e., multimedia content) to present the complex results of their investigative research in a comprehensible way to their readers. Additionally, the utilization of media types, text styles, and visual representations by Austrian journalists, when they tell their investigative stories, is analyzed.

II. INVESTIGATIVE JOURNALISM

Investigative journalism is understood as a subgenre of journalism which is particularly concerned with the uncovering of grievances and consequently with a social balance between officials and decision-makers as well as administrative bodies, private companies and the mass of citizens [4]. Historically, early investigative contributions therefore dealt with social differences between wealthy classes and grievances among, for example, the working class and other socially disadvantaged groups. Often journalists actively represented political positions personally and through their investigative work in order to stand up for social justice [5] [6].

In the course of time, several subject areas emerged from this motif which became classic content for investigative journalism: abuse of authority, institutional mismanagement, self-enrichment and corruption, nepotism or selfish acts at the expense of the general public [7]. However, since such grievances do not only affect the actions of individuals – irrespective of whether they are politicians, office holders or private (economic) power holders – but also gain social reach and relevance through the position of the actors, they are not just individual incidents. Instead, investigative stories can deal with structural grievances and systemic injustices that arise from the misconduct of individuals or institutions. However, this may also mean effects of private or political practices on society, on the environment, or on security [8]. The subject areas of investigative journalism have broadened with society over the course of time, but they have in common the social relevance and uncovering of grievances and misconduct.

Due to their social relevance, however, affected politicians or power holders have little interest in the publication of investigative contributions since this shows behavior that is not socially desirable or is even punishable. The task of investigative journalists is to obtain and publish this information against resistance [7]. Due to these sometimes adverse circumstances, researches in the past were often carried out under specification of false facts, identities or concealment of the journalistic intention. Today this approach is legally questionable and has therefore become rarer [6].

At the same time, it is also possible to search for interlocutors and informants who provide anecdotes, facts or documents. However, the role of informants is particularly critical when the amount of information provided by informants exceeds the information produced by the journalists themselves during the creation of a contribution. Journalists can also be manipulated in this way. After all, the level of information is relevant for the overall picture of a story. For example, information provided must be regularly checked for the intention of the informant in order to prevent journalists with publications from merely being in the interest of an informant and not in the public interest [5] [7].

Investigative journalism has a tradition in Austria back to the 19th century. However, due to the social upheavals during the World Wars, this type of journalism disappeared for a time; only from the 1970s did a culture for it emerge again [9]. On the one hand, this was due to social change, but in the course of the next decades magazine foundations also contributed to the fact that individual media – mostly weekly newspapers such as *Falter*, *News* and *Profil* – gained a reputation as investigative media [10]. Digitisation also had an impact on the sector, and the online platforms *Addendum* and *Dossier* were founded with a focus on enigmatic/ investigative journalism.

III. Online journalism and forms of presentation

Online journalism is characterized by preparation and presentation [11]. Especially in its initial phase, the Internet was perceived as an additional presentation platform for journalistic content. In terms of content the contributions of all previous media types can be adopted, whereby we can exploit the fact that online platforms have de facto no temporal restrictions and take over an archive function for print, TV and radio. For a long time, this archive function was seen as the main use of the Internet in journalism. However, the possibility of permanently updating the content is also a key aspect. Ideally, care should be taken during implementation to ensure that content is up-to-date, but updates can also provide great added value in some cases [12].

Real online journalism is characterized by interactivity and multimedia. Originally interactivity referred to the possibility of interconnecting different websites, but the possibilities of multimedia implementations developed rapidly, since various modes of media are embedded and presented [13].

Due to the different possibilities of the forms of presentation in online media, online journalism has to be thought of in a more diverse way. The publication of content is no longer limited to the predetermined medium, but can be adapted to different contents [14]. Although the implementation of multimedia content requires more preparation and thus represents a certain barrier to frequent use, the change must start with the journalistic concept. Journalists have to consider different forms of presentation during the research and produce the respective material [15].

The first online presentations primarily used text. In the case of longer contributions, this often has a negative effect on the capacity of readers to understand [16]. In order to reduce this deficit, text alone can create visual stylistic breaks using structure or highlighted elements and set priorities or separate content in extra columns or highlighted text elements in info boxes – just like in a printed presentation. The emphasized elements can focus on information, but also offer space for summaries or further information. Texts can also be loosened up by images, which, depending on the format, set different priorities and can push the attention of readers in certain directions, both visually and in terms of content. The advantage of images is not only their visual variety, but they can also illustrate content and put readers into a scene through atmospheric elements [16]. According to previous studies many users react strongly to visual stimuli. Thus, media consumption can be increased by images. Readers remember them better than text and they can remind content more strongly through optical support [17].

Graphics and illustrations also help to understand data, processes, or interrelationships. Depending on the content, various display formats are available to present statistical correlations, geographical, temporal, or content-related relationships. Readers are thus not only dependent on text to understand complicated content, but the visual implementation of sequences or processes makes it easier to understand them. Various visualizations are available for different data formats, depending on content and context [18]. However, it is relevant that graphics require clear and comprehensible labels, so that users can actually grasp this content and do not fail to understand the visual representation [16].

To bring content to life, videos and audios can be used to convey scenic content with atmospheric background information to readers. These should activate users, whereby the use is extended to different communication modes and readers themselves are actively involved in the use [16].

Nevertheless, readers benefit most when content is implemented in such a way that they can experience it for themselves, leaving the speed of perception and selection to the readers themselves – i.e., when it is implemented interactively. In this way, interactive information graphics or data formats can provide additional dimensions. This increases the entertainment factor during reception, which allows users to actively receive and remember more content. Timelines or information graphics can combine different media formats, use vertical or horizontal navigation individually and make processes comprehensible in detail with the help of additional information. This allows users to grasp parallel strands of action or effects between certain events. By using interactivity, these events can be experienced at second hand, so that they remain in mind.

Data processing can also represent different dependencies through interactivity, which makes abstract data connections an individual experience. The same applies to networks or maps that can be moved out of their two-dimensional representations through interactivity and, for example, receive an additional temporal component, so that users can not only perceive geographical connections or personal interrelationships in networks, but can also experience independently how these developed in the course of time.

IV. Themes in investigative journalism

Based on these findings, the websites of Austrian media – e.g., online media, magazines, weekly newspapers – were not examined on a daily basis since online investigative content does not occur to a high extent. In accordance with this starting point, four Austrian media dedicated to investigative journalism – "Addendum", "Falter", "Dossier", "News" and "Profil" – were analyzed with regard to the forms of presentation used for investigative contributions. The aim was to analyze the content of ten different investigative journalistic online articles per media, which were finally collected from all media on 15 July 2018. The selection of the investigative contributions was based on the theoretically derived categories: mismanagement, abuse of authority, self-enrichment, nepotism, corruption, assertion of self-interests at the expense of the general public, fraud/ ignorance/ conflicts of interest, systemic grievances/ injustice, ecological grievances, security risks, transparency [3].

The medium *News* had to be removed from the analysis, since at the time of the study no article with investigative content was available on the medium's website – although the weekly paper is characterized by such content both historically and to some extent in current print editions.

53 evaluations of investigative content were recorded in individual articles (Table I) [3]. Overall, transparency was the most pronounced category with eleven entries. One of the reasons for this could be the strong official secrecy in Austria. Nevertheless, the so-called traditional theme categories, as proposed by Ludwig, represent a majority of the themes [7].

It is also evident how individual topics are set. Three out of four media have a broad range of topics and achieve at least one entry in two thirds of the categories. All in all, it can be seen from the topics of the analyzed media that they are widely spread. Only the topic *security risks* is not dealt with in the sample of articles that has been analyzed.

The high number of contributions on the subject *transparency* suggests that the way in which topics are set has changed and that investigative journalists perceive their role differently. More articles are published, for example, which independently analyze a topic, its functions and interrelationships and make it publicly available without external cause. Thus, the public sector is controlled at an own initiative, often with data-driven content. Depending on the medium, specific priorities are set. For example, *Dossier* already has a strong focus on advertising expenditures of public institutions from its founding history, and expenditures for funding are also suitable for data-related contributions.

TABLE I. INVESTIGATIVE TOPICS

Investigative contents	**Naming of media**			
	Addendum	***Falter***	***Dossier***	***Profil***
Mismanagement	0	1	4	2
Abuse of authority	2	3	2	0
Self-enrichment	1	0	2	0
Nepotism / cronyism	1	2	1	0
Corruption / bribery	0	0	0	1
Self-interest at the expense of the general public	2	3	1	3
Fraud, ignorance, conflicts of interest	1	1	2	1
Systemic grievances	2	1	0	2
Ecological grievances	0	1	0	0
Security risks	0	0	0	0
Transparency	3	1	4	3

Although transparency is the most common topic, the further distribution shows that the "classical" topics as defined by Ludwig together make up the majority of contributions. The canon of topics that constitutes the core of investigative journalism has not changed overall, but has been expanded [7].

V. Implementation of forms of representation in investigative journalism

Based on the literature and various practical examples, a category catalogue was created for the forms of presentation. These are the categories used in the analysis within which a distinction was made between static and interactive formats:

- Highlighted quote
- Infobox
- Header image
- Full image
- Picture in continuous text
- Provided picture
- Audio embedding
- Video
- Information graphic
- Scale
- Bar/bar chart
- Relative bar chart
- Dot graphic
- Scatterplot/ connected point graphic
- Linear Graphics
- Pie chart
- Treemap/ Area Chart
- Radar Graphics/ Network
- Arc graphics
- Flow chart

- Timeline
- Map
- Original document
- Interactive info box
- Picture galleries
- Document Galleries
- Vertical Timeline
- Horizontal Timeline
- Interactive information graphic
- Interactive bar/bar chart
- Interactive Relative Bar Chart
- Interactive point graphics
- Interactive scatterplot/ linked point graphic
- Interactive linear graphics
- Interactive pie chart
- Interactive treemap/ area chart
- Interactive radar graphics
- Interactive arc graphics
- Interactive Flowchart
- Interactive network
- Interactive map
- Scrollytelling

The analysis revealed that the use of forms of representation is strongly related to the medium itself and possibly also to its founding history (Table II). For example, the print media *Falter* and *Profil* hardly use any different forms of presentation. The two platforms *Addendum* and *Dossier*, which were founded as online media, use different forms of presentation – especially with a focus on data visualization.

Common to *Addendum*, *Falter* and *Profil* is the default use of header images. *Addendum* inserts them over the entire width of the page, *Falter* and *Profil* use a preceding image that is located at the beginning of the article. These images are used in combination with headlines and teaser texts. The arrangement around the header image differs for all three media. *Addendum* inserts title and teaser in the image. *Falter* positions the title and teaser above the image. *Profil* inserts the title above the image, but the teaser below the image. However, in the evaluation the category "Header Image" is identified only 28 times. This is due to the fact that two videos that had been leaked to *Falter* have been inserted instead of the header image. Apart from these preceding images and videos *Falter* does not use any other forms of representation. *Profil* uses only a few design options: once a source document is inserted into the continuous text, as well as two images were inserted to loosen up the text. *Addendum* and *Dossier*, that have been founded as online media with a focus on investigative and enigmatic journalism, use different forms of presentation. They show pictures for illustration and highlighted quotes to change the reading flow, as well as original documents to substantiate statements and findings. Also noticeable is a focus on data collection and data processing, which are used very frequently in connection with transparency-related content.

TABLE II. ANALYSIS OF THE FORMS OF PRESENTATION USED

presentation mode	**Naming of media**			
	Addendum	***Falter***	***Dossier***	***Profil***
Highlighted quote	20	0	14	0
Infobox	4	0	0	0
Header image	10	8	0	10
Picture in continuous text	9	0	2	2
Added image	0	0	2	0
Audio embedding	6	0	0	0
Video	5	2	1	0
Information graphic	1	0	0	0
Scale	3	0	0	0
Bar/bar chart	13	0	1	0
Relative bar chart	0	0	2	0
Linear Graphics	0	0	3	0
Radar Graphics/ Network	0	0	2	0
Map	1	0	0	0
Original document	1	0	9	1
Interactive info box	4	0	0	0
Document Galleries	7	0	0	0
Interactive bar/bar chart	3	0	0	0
Interactive linear graphics	2	0	0	0
Interactive map	1	0	0	0

Although they appear in the list for the sake of completeness, the audio contributions of *Addendum* may not be counted as separate forms of presentation. They do not represent any added value of their own, as they are merely a soundtrack to the contributions, which should allow readers a further mode of consumption (keyword: accessibility). Since the audio tracks of the contributions appear with different lead times, not all contributions included an audio track already, and consequently only six audio files appear in the analysis.

The use of data visualizations differs with *Addendum* and *Dossier* since *Addendum* uses much more interactive elements, but in principle the classic visualization using bar charts and linear graphics is similar. However, it must be added that the bar and column diagrams and linear diagrams in *Addendum* are only interactive in the sense that the respective data set is displayed during a mouse-over. Since both media use the Datawrapper program for such data visualizations, the difference between the data visualizations of *Dossier* and *Addendum* could be due to the different financial resources of the two media.

Addendum has the most different forms of presentation with 90 elements in 16 categories, followed by 36 forms of presentation in nine categories that could be identified in *Dossier*. In principle, it can be stated that highlighted quotations are most frequently used (20 respectively 14 times) to loosen up long texts and to emphasize individual passages.

Addendum uses additional forms of presentation in all contributions; not a single analyzed contribution contains only a header image. Nevertheless, interactive forms of representation are relatively rare: only four different forms of representation and 13 elements have been identified. The other contributions use between two and 21 elements, usually the contain between two and five forms of presentation. Depending on the content of the contribution, different forms of presentation such as tables, diagrams, maps and graphics are used. As a minimum, header image and audio – which, as mentioned above, offers no added value in terms of content – are used as illustrative forms of presentation There are only two contributions that are based on text only, apart from the header image and the audio file.

The ten contributions of the platform *Dossier* contain a total of nine different forms of presentation, including quotations, images, videos, original documents and data pieces (bar/bar chart; ratio-oriented bar chart, linear chart, network).

The most frequently used form for loosening up texts are embedded quotations, which were used a total of 14 times in five contributions. In second place follow pictures that are embedded in continuous text, as well as provided pictures. Altogether eleven pictures have been identified in six contributions. A video has also been used to illustrate and display an advertising insert, but in fact it has been used for purely illustrative purposes. Apart from the illustration, the video does not offer any added value in terms of content. In the case of *Dossier*, it is noticeable that data formats in particular are used in the sense of time series – a visual focus that can be explained by the focus on data journalism in terms of content.

In general, it can be stated that *Dossier* selects display formats on a content-oriented basis, so the display formats also vary depending on the content. It is noticeable that quotations are often used to loosen up or emphasize content. Pictures and illustrations are used according to individual needs. However, in this way individual contributions are created which hardly contain an optical loosening of the continuous text. On the other hand, various and different forms of presentation are used in individual contributions to support and illustrate content.

VI. Usage of Visual Presentations by Austrian Online Journalists

To complement the content analysis of the online presence of four Austrian media that are well known for publishing investigative contributions, we conducted a study among Austrian journalists to identify how they actually use different journalistic text styles, media elements and visual representations to tell their investigative stories. A quantitative approach based on an online survey was chosen to determine the opinion of journalists engaged in investigative journalism. Since only a small number of journalists is active in this specific type of journalism the survey was restricted only to journalists that work for the journalistic media types that are likely to include investigative contributions: daily newspaper, weekly newspaper, magazine, TV, radio, and Internet.

The online survey was conducted in June 2018. 2716 Austrian journalists have been invited per email to participate in the online survey. Since the email addresses have been derived from the Austrian media handbook (*Medienhandbuch*) the survey type is *Internet survey of specifically named persons* [19]. 250 persons accessed the questionnaire, but only 166 respondents actually started the questionnaire. There was a significant number of respondents that broke off after the first page. This might be due to the fact that those respondents felt involved by the email in the initial approach, but realized lateron that their working practice did not correspond to a more precise definition of investigative journalism. A total of 109 respondents completed the questionnaire, however some of them did not fill in all questions (i.e., partial interview). The overall response rate is only 4.01 % (based on the number of eligible contacts) [19]. The low response rate is influenced by the fact that only a small number of journalists are active in investigative journalism and only those journalists of selected media types have been addressed.

The majority of the responding journalists is male: 65% of male respondents vs. 35% female respondents. Journalists working on investigative topics are typically somewhat older (Table III).

TABLE III. Age distribution (N=107)

Age	Percentage
18 to 30	11.2 %
31 to 40	24.3 %
41 to 50	23.4 %
50 plus	41.1 %

Most journalists work on investigative stories on politics (40), economy (37), or society & social issues (31). Stories on sports (14), health (11), or science (11) are less frequent. Stories on culture are very rare (4). Other topics including crime, local news, and justice have been mentioned 22 times (Fig. 1; N=84, multiple references possible).

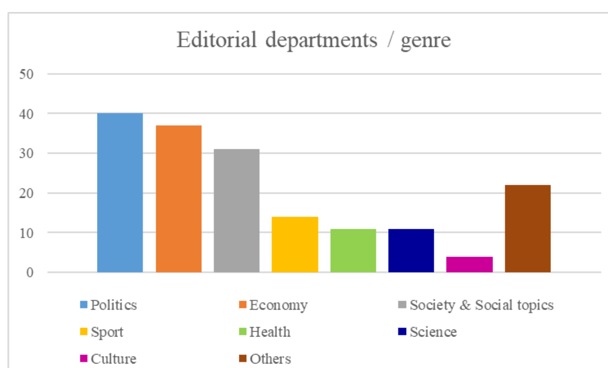


Fig. 1. Genres of investigative contributions (N=84, multiple references)

Profound investigative research is a time consuming and costly task. 18.5 % of investigative research lasts less than a day, but the largest share of 45.7 % corresponds to research that lasts up to one week. 30.9 % of investigative research takes up to one month and 4.9 % takes even longer (Table IV).

TABLE IV. DURATION OF RESEARCH (N=81)

Duration	Value	
	Absolute	*Percentage*
One day	15	18.5 %
Up to one week	37	45.7 %
Up to one month	25	30.9 %
More than a month	4	4.9 %

Journalists use different sources for their investigative research. Most journalists use information provided by reliable informants from their own network (86.6 %). 61 % use data from public web sources and an own on-site inspection. 59.8 % of the journalists get their information from publicly available documents while 43.9 % use secret documents. 52.4 % rely on transparent interviewees that may be mentioned, and 46.3 % use transparent interviewees that have to be made anonymous. Data journalistic analysis is performed only by 37.8 % of the journalists (Fig. 2; N=80, multiple references possible).

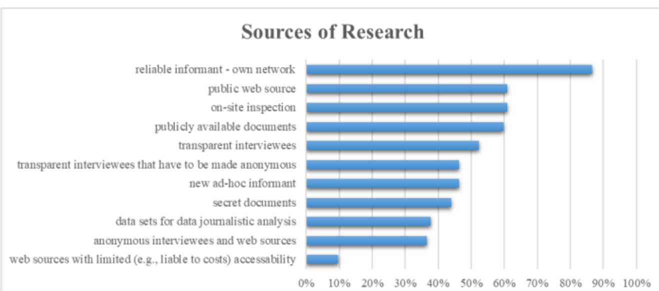


Fig. 2. Sources of investigative research (N=80)

Journalists use various forms of (visual) presentation of their investigative content. Text is the dominant form of presentation. 85.7 % of the respondents (N=84) use text as the primary form of presentation. 64.3 % use pictures and 50 % use information graphics. Online media allow journalists to incorporate video and audio in their investigative contributions as well, however they are used much less (video: 14.3 %, audio: 9.5 %).

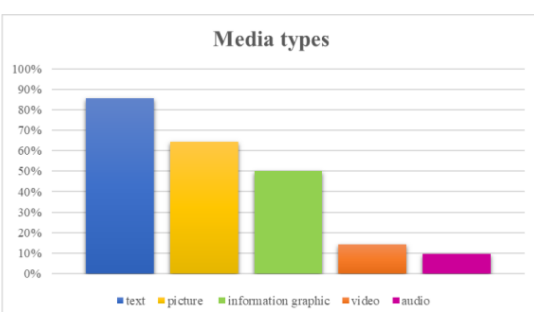


Fig. 3. Media types used by journalists for investigative content (N=84)

Additionally, journalists working for daily newspapers, weekly newspapers, and magazines (N=47) were asked how important they consider the different media types for the representation of investigative content. Remark: Since journalists working for TV and radio are much more limited in their choice of (visual) representation their answers are excluded in the following results. For this purpose a four-point Likert scale has been applied: very important | rather important | less important | not important (Fig. 4).

Similar to the actual usage of the analyzed forms of presentation (i.e., media types) the results show that text is most important. All journalists consider text to be important: 91.3 % very important or 8.7 % rather important. Pictures are important as well for 76.1 % (very important: 45.7 % or rather important: 30.4 %). Only 23.9 % rate pictures less important for investigative content. The results of the survey show a quite similar outcome for the type information graphics. 78.3 % of the journalists rate infographics to be important (very important: 41.3 % or rather important: 37 %). While 17.4 % expect infographics to be less important, a small number of 4.3 % consider them not important. Video is important for 60.9 % of the journalists (very important: 17.4 % or rather important: 43.5 %), while audio is important for 45.7 % of the journalists (very important: 13.1 % or rather important: 32.6 %).

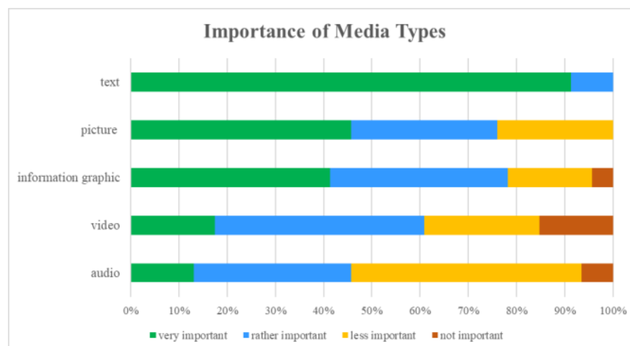


Fig. 4. Importance of media types (N=47)

Since journalists consider text as the most important form of representation for investigative content we analyzed which text styles the journalists use most frequently (based on the usage in daily newspapers, weekly newspapers, and magazines; Fig. 5). The most common text styles are report (95.3 %: 83.7 % very frequent, 11.6 % frequent) and reportage (69.8 %: 30.3 % very frequent, 39.5 % frequent). Less popular, but still frequently used are news items (55.8 %: 23.3 % very frequent, 32.5 % frequent), column (46.5 %: 20.9 % very frequent, 25.6 % frequent), editorial comment (39.5 %: 13.9 % very frequent, 25.6 % frequent), and portrait (48.8 %: 11.6 % very frequent, 37.2 % frequent).

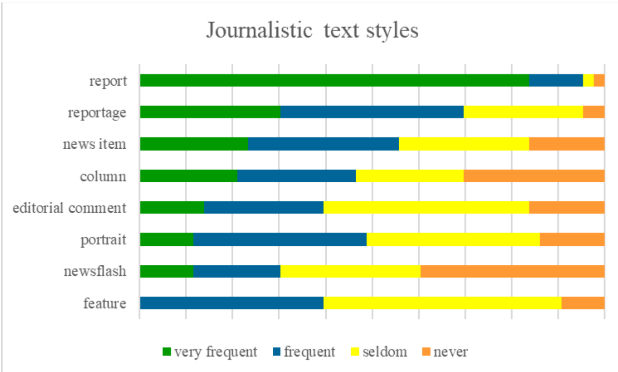


Fig. 5. Usage of journalistic text styles for investigative content (N=43)

VII. UPSHOT AND PROSPECTS

Traditionally, in scientific literature the contents of investigative journalism have been limited to mismanagement, abuse of authority, self-enrichment or corruption – contents that uncover misconduct by officials, dignitaries, private companies or those in power in society. Investigative journalism therefore sees itself as a social control authority. Instead of reporting on events, an attempt is made to point out relevant events and to direct a social and media focus on stories that otherwise happen outside public perception. In addition to content on active events or actions investigative journalism also examines systemic grievances that can lead to injustice without human intervention. Furthermore its task is to make non-public content public in order to prevent abuse of power and increase the level of public information in the sense of transparency.

The four analyzed media focus on showing non-public content and thus transparency-related stories. However, articles on "classical" categories of investigative journalism as defined, for example, by Ludwig [8] predominate and thus content on abuse of authority and power, mismanagement or the enforcement of self-interests. Due to the manifold relevance of content, 53 investigative aspects could be identified in 40 contributions.

Different visual and interactive elements can be used to make it easier for readers to understand investigative content. They also increase attention at reception and the likelihood that users can actively remember content. Despite the variety of static and interactive forms of presentation, however, it became apparent that media that developed their online presence as a second communication channel (*Falter*, *Profil*) have problems in using these possibilities. Only in exceptional cases, videos or documents are embedded and active attention is paid to loosening up the text flow. At the same time, potential is also actively dispensed with, since, for example, info boxes are only used in the print edition but are not implemented online. Purely online media use many more elements, including pictures for illustration and highlighted quotations, to change the reading flow, as well as original documents to substantiate statements and findings. It is also striking that the focus is on data formats which are used particularly frequently in connection with transparency-related content. Although the use of data visualizations differs between *Addendum* and *Dossier* due to the interactivity of *Addendum*, the basic visualization by bar charts and linear graphics is identical. The experimental factor, which is made possible by the complete implementation of interactive data formats, is also not exploited fully by *Addendum*, although different forms of presentation are used much more frequently for *Addendum* than for *Dossier*.

The study among Austrian investigative journalists to identify how they actually use different journalistic text styles, media elements and visual representations to tell their investigative stories revealed that text is the dominant form of presentation (86 %) actually used in investigative reports. Pictures (64 %) and information graphics (50 %) are frequently used as well. However, video (14 %) and audio (9 %) are seldom used to present investigative content. When journalists had been asked which media types they consider being important for the representation of investigative content we got similar results. All journalists consider text to be most important (100 %), followed by information graphics (78 %) and pictures (76 %). Video is important for 61 % of the journalists, while audio is important for 46 % of the journalists. The focus on text may be due to the fact that investigative journalists still heavily rely on classic sources of information. For example, data journalistic analysis is performed only by 37.8 % of the journalists in our survey. Consequently, the most common text styles are report (95 %) and reportage (70 %), followed by news items (56 %), column (46 %), editorial comment (39 %), and portrait (49 %).

Overall, the following conclusions result for the visual representation of investigative online journalism in Austria:

- Investigation is sparsely represented in Austria, at least in the more broadly effective area of online journalism – despite its corresponding social relevance.
- Media that are primarily geared to online presence use more and more different visual and structural forms of presentation to make content easier to understand. However, the interactivity factor was only used to a limited extent with added value, at least during the survey period. And: These media have comparatively low audience reach, so many contents do not reach a large part of the population.
- Within the media focused on the online presence, there is a difference in the presentation which can at least partly be traced back to the different resources.

All in all, there is a clear need for Austrian media to catch up in investigative online journalism in general, and within this, in turn, there are major omissions – or in other words: potentials – in the field of visual representation of it. Like everywhere else in journalism today, the question of financing is decisive for the quality of implementation, even among those media that are currently making better progress. A solution to this problem could be reproducible prototypes and tool-based processes for the representation of investigative online journalism, which ensure a more resource-saving implementation and thus enable the media to ultimately make highly relevant social content available to a larger number of people more simply and cost-effectively.

Acknowledgment

The online survey has been conducted with the help of students in the Bachelor program Information, Media and Communication: thanks to Silvia Gassner, Manuela Illichmann, Christine Madner, Christian Priedl, Denise Reinprecht, Sarah Reinprecht, Nathalie Sched, Maximilian Unterrieder, and Lisa Willisits.

References

[1] J. Ludwig, Investigatives Recherchieren. Praktischer Journalismus, Band 48. Köln: Halem, 2017.

[2] A. Cavar, “Wir haben uns rücksichtslos kannibalisiert“. Internet: http://updatedigital.at/news/medien/tag-des-qualitaetsjournalismus-zur-finanzierung-von-qualitaetsjournalismus/68.129, Apr. 30, 2013.

[3] S. Braunisch, “Multiperspektivische Darstellungsformen im investigativen Onlinejournalismus am Beispiel österreichischer Medien“. Master Thesis, University of Applied Sciences Burgenland, Austria, 2018.

[4] S. Nazakat, “Investigative Journalism Manual”. Konrad-Adenauer-Stiftung. Internet: http://www.investigative-manual.org/wp-content/uploads/2016/09/20171026-IJM_final.pdf, 2010.

[5] M. Haller, Recherchieren. (6th edition). Konstanz: UVK Verlagsgesellschaft, 2004.

[6] D. Lorenz, Journalismus. Stuttgart, Weimar: J.B. Metzler, 2002.

[7] J. Ludwig, Investigativer Journalismus: Recherchestrategien – Quellen – Informanten. Konstanz: UVK Verlagsgesellschaft, 2002.

[8] R. Rodriguez, A Vital Responsibility In Need Of Support. Nieman Reports, vol. 62, no. 1, Cambridge: Nieman Foundation, 2008, pp. 62-65.

[9] H. Fabris, Der verspätete Aufstieg des Journalismus in der zweiten Republik. In: H. Fabris and F. Hausjell (eds.), Die Vierte Macht. Zur Geschichte und Kultur des Journalismus in Österreich seit 1945, Wien: Verlag für Gesellschaftskritik, 1991, pp. 11-28.

[10] A. Schartmüller, Investigativer (Print-)Journalismus in Österreich: Wer betreibt ihn (noch) und wie funktioniert er? Wien: Universität Wien, 2009.

[11] J. Weiss, Das Internet und die klassischen Medien. Konvergenz – Konkurrenz oder Komplementierung? Frankfurt am Main: Peter Lang. Europäischer Verlag der Wissenschaften, 2003.

[12] A. Godulla, and C. Wolf, Digitale Langformen im Journalismus und Corporate Publishing. Wiesbaden: Springer Fachmedien, 2017.

[13] P. Schumacher, Rezeption als Interaktion. Wahrnehmung und Nutzung multimedialer Darstellungsformen im Online-Journalismus. Baden-Baden: Nomos, Edition Fischer, 2009.

[14] J. Radü, Technologie als Chance. Auf welche Weise Smartphones, Tablets und die Medientechnologie der Zukunft journalistische Qualität sichern helfen. In: L. Kramp, L. Novy, D. Ballwieser, and K. Wenzlaff. Journalismus in der digitalen Moderne. Einsichten . Ansichten – Aussichten. Wiesbaden: Springer Fachmedien, 2013, pp. 173–184.

[15] C. Jakubetz, Universalcode 2020. Content + Kontext + Endgerät. Konstanz: UVK Verlagsgesellschaft, 2016.

[16] S. Sturm, Digitales Storytelling. Eine Einführung in neue Formen des Storytellings. Wiesbaden: Springer Fachmedien, 2013.

[17] A. Godulla, and C. Wolf, Die Usability neuer Darstellungsformen im digitalen Journalismus. In: G. Hoofacker & C. Wolf (eds.), Technische Innovationen – Medieninnovationen? Herausforderungen für Kommunikatoren, Konzepte und Nutzerforschung. Wiesbaden: Springer Fachmedien, 2016, pp. 62-75.

[18] A. Smith, C. Campbell, I. Bott, G. Parrish, B. Ehrenberg, et al., Visual Vocabulary. Financial Times. Internet: https://github.com/ft-interactive/chart-doctor/tree/master/visual-vocabulary, 2016.

[19] The American Association for Public Opinion Research, Standard Definitions - Final Dispositions of Case Codes and Outcome Rates for Surveys. April 2015.

Modeling User Interface Adaptation for Customer-Experience Optimization

Christian Märtin, Christian Herdin, and Bärbel Bissinger
Augsburg University of Applied Sciences
Faculty of Computer Science
Augsburg, Germany
{Christian.Maertin, Christian.Herdin}@hs-augsburg.de; baerbelbissinger@web.de

***Abstract*—The customer journey in digital marketing defines several touch points, where interested users can directly interact with an e-business platform. In order to convert a user into a buyer, persona-based a priori adaptations of the user interface can be combined with dynamic adaptations at runtime with the goal to optimize individual customer experience and guide task accomplishment. This paper examines customer experience optimization for scenarios from a cosmetics industry e-business portal with the SitAdapt 2.0 system. Dynamic adaptations are triggered by situation rules based on the continuous analysis of the users' varying cognitive and emotional situations during a session. The model-based adaptation process exploits models and patterns for the rapid generation of user interface modifications.**

Keywords—customer journey; user experience; customer experience; situation analytics; situation rules; emotion recognition; eye-tracking; HCI-patterns

I. Introduction and Related Work

Digitalization in marketing can be seen as a straightforward approach to designing and implementing IT-based solutions for the generic steps of the customer journey. A customer journey is a customer's interaction at several touch points with a service or several services of one or more service providers in order to achieve a specific goal [9]. More focused on purchasing a product, the customer journey can be defined as an iterative process that includes touch point based interactions with a provider or a business during a pre-purchase, a purchase, and a post-purchase phase [13]. The journey could include experiences from earlier purchases and affect future purchases. In this view no fixed a priori purchase goal is necessary, but the service provider would try to arouse the interest of potential customers in the pre-purchase phase. At all touch points between the provider and the customer, one has to distinguish between the customer view and the provider view. It must be the provider's goal at every touch point, to create a situation that leads to optimum user experience (UX) for the potential customer.

UX during the customer journey is often described as customer experience. As an extract and synthesis of earlier research efforts customer experience can be seen as "a multidimensional construct" that focuses on "a customer's cognitive, emotional, behavioral, sensorial, and social" reactions to the offerings of a provider or a business "during the customer's entire purchase journey" [13].

With SitAdapt [14], [15] we have developed a software architecture for situation analytics and for integrating adaptive behavior into web- or app-based interactive applications. SitAdapt fulfills the requirements for automating essential parts of the customer experience optimization process as well as for various other domains from medical monitoring to driver assistance systems. Possible adaptations are modeled within the PaMGIS MBUID framework [5], [6]. They are triggered by situation rules and generated by activating and exploiting domain-dependent and independent HCI-patterns. In this paper we present our preliminary lab-based results for using the current implementation SitAdapt 2.0 with a new rule editor and an advanced situation interpreter within the e-commerce domain[1].

The paper includes the following main contributions:

- Discussion of a new model-based approach [17] for automating customer experience optimization
- Defining the potential for software adaptation [24], [12], [19], [20] based on situation analytics [3], context-awareness [22], and situation-awareness [7]
- Demonstrating the suitability of emotion recognition and bio-signal tracking for triggering user interface modifications [8], [19], [21], [23].
- Detailing the adaptation process and workflow for the e-business domain

The remainder of the paper is structured as follows:

Chapter II introduces the SitAdapt 2.0 system with its new rule editor. Chapter III first introduces possible adaptation features and defines example scenarios for generic and individual situations that are occurring in different phases of the customer journey when visiting a cosmetics business portal. Some of the possible SitAdapt 2.0 use-cases are demonstrated. After this, the chapter discusses the modeling and generation of adaptations. Chapter IV concludes the paper.

[1] Part of this work was carried out in cooperation with Dr. Grandel GmbH, Augsburg, Germany. We greatly acknowledge the opportunity to run the SitAdapt 2.0 tools and user tests on their enterprise e-business platform.

II. SitAdapt 2.0

The SitAdapt 2.0 runtime environment is integrated into the PaMGIS (Pattern-based Modeling and Generation of Interactive Systems) development framework. The framework allows for modeling and generating responsive behavior in the user interface and has now been enhanced towards dynamic adaptation by situation interpretation at runtime.

The architecture (Fig. 1) consists of the following parts:

- The *data interfaces* from the different devices (Tobii eye-tracker[2], Empatica wristband[3], Noldus Facereader[4], metadata from the application)

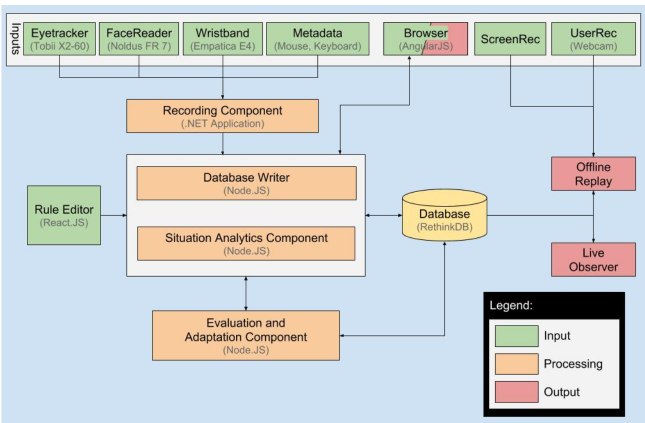


Fig. 1. SitAdapt 2.0 Architecture

- The *recording component* synchronizes the different input records with a timestamp, records the eye- and gaze-tracking signal of the user and tracks the emotional video facial expression as a combination of the six basic emotions (happy, sad, scared, disgusted, surprised, and angry) based on Ekman's model [4]. Other recorded data about the user are, e.g., age-range and gender [15]. The stress-level and other biometric data are recorded in real-time by a wristband. In addition, mouse movements and keyboard logs are protocolled [11].

- The *database writer* stores the data from the recording component and from the browser in the database in the form of discrete raw situations and manages the communication with the rule editor. Raw situations are generated at each tick of a predefined time frame varying from 1/60s to 1s.

- The *rule editor* allows the definition and modification of situation rules, e.g. for specifying the different user states (e.g. if a happy state is observed, it will only become relevant, if the state lasts more than five seconds and the grade of the emotion surpasses a certain activation level). For experimenting with rule heuristics and observing users we built a prototypical web application for long distance travel booking. In addition we used results from our cosmetics industry user study [1] for finding plausible situation rules. Fig. 2 shows the creation of a simple situation rule with two conditions and one action. In this case only a dialog with the user is created. However, situation rules can also activate HCI-patterns in the PaMGIS pattern repository. These patterns are exploited at runtime to generate user interface adaptations from predefined UI-, task-, or domain-model fragments.

- The *situation analytics component* analyzes the sequences of raw situations with their parameters varying over time and condenses them to a situation profile holding the most significant information about the currently applying situations. Typical situations can be described in the form of situation patterns. The situation analytics component matches the raw sequences to such situation patterns. A set of typical domain-dependent and independent situation patterns is available in the PaMGIS pattern repository. Such situation patterns can serve as templates for creating situation rules with the rule editor, where an action part with one or more actions is added. New situation patterns can be discovered by running offline data mining tools, e.g., RapidMiner[5], on the raw situation sequences recorded during multiple sessions.

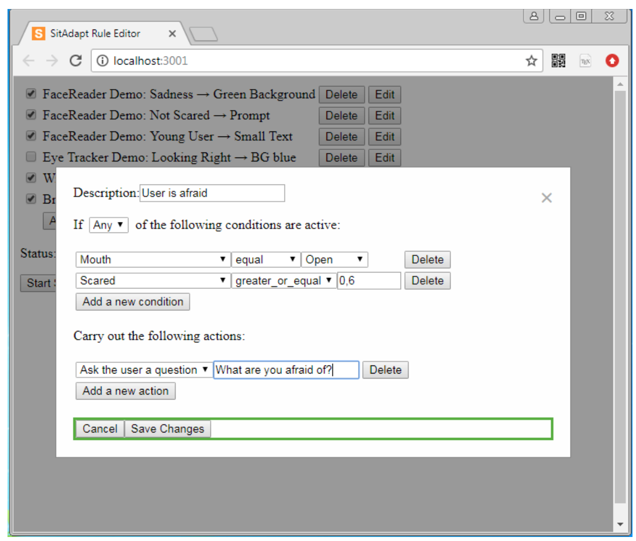


Fig. 2. SitAdapt 2.0 Rule Editor with an example rule exploiting visual emotions provided by Facereader

- The *evaluation and adaptation component* uses the situation profile provided by the situation analytics component to decide whether an adaptation of the user interface is meaningful and necessary at a specific moment. For this purpose the component evaluates given situation rules. Whether an adaptation is meaningful depends on the predefined purpose of the situation-aware target application. Goals to meet can range from successful marketing activities in e-business, e.g. having the user buying an item from the e-shop or letting her or him browse through the latest

[2] www.tobii.com
[3] https://www.empatica.com/en-eu/research/e4/
[4] www.noldus.com/human-behavior-research/products/facereader
[5] https://rapidminer.com

special offers, to improved customer experience levels, or to meeting user desires defined by the hidden mental states of the user. The adaptation component finally generates the necessary modifications of the interactive target application.

These architectural components are necessary for enabling the PaMGIS framework to support automated adaptive user interfaces. In the user interface construction process, the SitAdapt 2.0 evaluation and adaptation component cooperates with the models of the interactive application (abstract, concrete and final user interface model, context of use models, task and concept model) and can also access the HCI-patterns (not to be confused with the situation patterns) residing in the PaMGIS repositories to build the necessary modifications of the user interface at runtime.

III. Automating Customer Experience Optimization in E-Business

As a promising candidate domain for exploring situation analytics and situation-aware adaptation we have selected the e-business and e-commerce fields. In our current project we focus on a commercial cosmetics e-business portal.

A. Dynamic Adaptation Features

We have implemented dynamic adaptation features for pre-session, first session and recurring session adaptation. Typical adaptation features are related to the following areas:

Visual appearance of the application

- Gender or age specific coloring
- Gender or age specific image selections
- Soothing image or color selection
- Age specific element size
- Element ordering or widget selection dependent on age or emotional state
- Screen contrast dependent on clock time, bio-physical or emotional user state
- Font type, font size dependent on age, clock time, bio-physical or emotional user state

New user interface or content elements

- Tutorial-offering at first session or dependent of user age
- Help functionality, e.g. chat window, help menu item, tool tips, UI element tips dependent on user behavior
- Personalized fields and panes (user- and behavior-specific advertisement

Content-based adaptation

- Personalized product offers or suggestions
- Voucher offering dependent on user behavior
- User feedback functionality dependent on user behavior

B. Complex Situation Examples for a Cosmetics Portal

In a comprehensive user study with 9 female test persons we tested the usability, user experience and emotional behavior for several scenarios when interacting with a real-world cosmetics industry web-portal [1]. These tests served as the basis for finding domain-dependent situations and formulating situation rules. Due to the limited space we can only discuss some of the most interesting findings in this section.

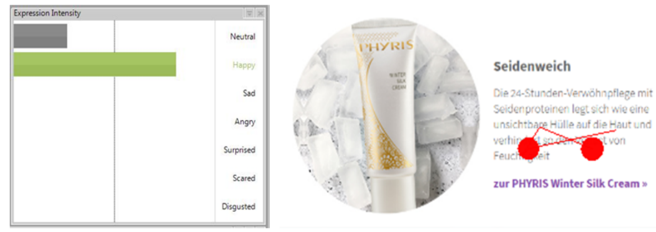


Fig. 3. Emotional reaction after finding the ideal product

In order to illustrate the potential of situation-aware adaptation we present some real-world situation examples and possible adaptive reactions. In the first example (Fig. 3) a test person is searching for a specific winter skin cream. Upon reading the detailed description of the product *Winter Silk Crème,* the user's emotional state significantly changes to *happy*. A situation rule could now exploit this knowledge to give additional information about other winter products. The improved customer experience near the purchase touch point can directly lead to a purchase of this and similar products.

In the next example (Fig. 4 and 5), the system has gathered a priori knowledge about the varying gaze behavior of test persons, who are known customers of the business or who are here for the first time, by distinguishing between the lab-created heat maps. The gaze behavior with respect to this image can be used to categorize anonymous users. The customer experience during the pre-purchase phase can be improved. When the system assumes a returning customer, the focus of her further customer journey will be put on showing aesthetic images, while in the other case more descriptive information will be given during the rest of the customer journey.

Fig. 4. Heatmap for customers of the business

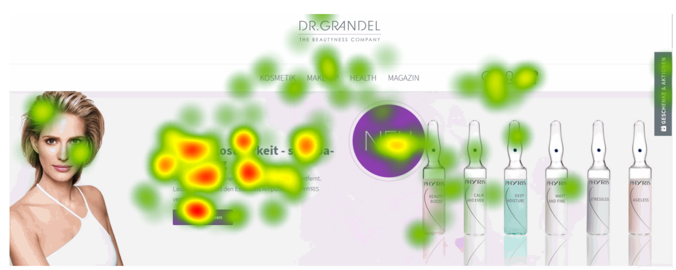


Fig. 5. Heatmap for first time users of the website

Another application area for using situation analytics in the e-business field is the evaluation and fine-tuning of pre-defined customer personas, which are used for pre-runtime adaptations and configurations of an application. Focusing on personas for a priori adaptation of the cosmetics portal can e.g. affect the visual appearance, the product content structure, the level of the product description language, the appearance of special advertisements, or the gaming and social media orientation of the website. Are test persons behaving like their respective personas or are there significant deviations from the expected behavior? This can be evaluated by comparing the situation profiles that come up during persona-adapted user tests with the typical situation profiles specified during the persona definition process. Vice-versa SitAdapt 2.0 can classify unknown customers or first time visitors into one of the given persona categories by analyzing the situations appearing during the session and by analyzing the users' behavior after situation-rule triggered adaptations.

All of these user observations and behavior evaluations as well as the adaptations of the interactive software are currently done in our situation analytics lab environment. The rapid evolution of visual and biophysical user tracking and monitoring technology will enable situation-aware individual adaptations for the end user in the near future.

C. Adaptation Modeling and Adaptation Process

By applying our MBUID approach, the modeling, generation and adaptation of the target website is done with the help of the PaMGIS framework and the integrated SitAdapt 2.0 system (Fig. 1). The PaMGIS framework is based on the Cameleon Reference Framework (CRF) [2], [18]. In the construction process first of all, the abstract user interface model (AUI) is generated from the information contained in the domain model of the application that includes a task model and the concept model (i.e. business model) and defines abstract user interface objects that are still independent of the context of use. This AUI model can then be transformed into a concrete user interface model (CUI) that already exploits the context of use model that includes the user, device, UI-Toolkit and environment model, and the dialog model, which is responsible for the dynamic user interface behavior. In the next step, the final user interface model (FUI) is generated by parsing the CUI model and by exploiting the context of use model and the layout model [15].

The first displayed version of the product e-commerce website is already adapted to the user. For example by using the age and target device information from the context of use model. The SitAdapt 2.0 system permanently monitors the user and recognizes the situations she or he is experiencing while viewing the webpage and interacting with the user interface. The evaluation component recognizes in the first example (Fig. 3), that the user reads the text attentively and that the level of the *happy* emotion surpasses a given minimum duration (e.g. more than 5 seconds). These data come from the raw situation sequences stored in the database by the recording component. The various inputs from the Facereader (emotion) and eye tracking (text screen field) and the metadata of the website (URL) were evaluated just in time by the situation analytics component that has created the following situation profile of the current situation:


```
<SituationProfile>
      <TargetApplication>Desktop_PC
      ...
      <Situation_product_view>
            <FUI_link> product_view
            <AUI_link>
            model_AUI_product_view_1
            <CUI_link>
      model_CUI_product_view_1
            <Dialog_link> product_view
            <Task_link> product_view
            <Concept_link>
      model_concept_Product_view_1
            <Eye_Tracking>
      Product_Textbox_Product_1(10 s)
            <USRUA_Age_Range>30-50
            <USRUA_Gender>female
            <EmotionalState> happy
            <UserPsychologicalState>
            <BiometricState>
                  <Pulse> normal
                  <StressLevel> green
                  ...
      </Situation_product_view>
</SituationProfile>
```


The evaluation and adaptation component examines the situation profile to decide, if an adaptation can take place. This is usually achieved by activating the responsible sub-set of situation rules in the rule editor (Fig 2.). Alternatively, the programmer or web designer can directly provide code for interpreting the situation profile in the web application client or server, which is triggered when the user interacts with specific elements of the user interface

In the concrete example, the situation rules specify that additional information about other winter cosmetic products should be displayed in this particular situation. The decision can be refined by also taking into account the user persona, if it is already known. For a strictly goal-oriented persona, a new window with additional product information may be shown. For a more cautious persona, a question text may appear, whether additional product information about winter products is welcome.

The evaluation and adaptation component now starts the adaptation process, which leads to the generation of a modified user interface. A new CUI and subsequently a FUI is generated and displayed to the user. The PaMGIS modeling environment must provide all the necessary models, model variants and model fragments necessary for user interface modifications. User interface models may contain links to HCI-patterns that can facilitate user interface code generation. More complex adaptations may also activate different tasks specified in the task model and require the activation of non-UI service code.

By observing the users' emotional behavior after such adaptations, the quality of the situation rules and the respective adaptations can be evaluated and rated. Such information can be used offline for refining the situation rule set for later use.

IV. Conclusion

SitAdapt 2.0 is an advanced architecture for automating user interface adaptation for responsive web-applications that were constructed with the PaMGIS MBUID framework. This paper has demonstrated the flexibility and versatility of this new approach by testing it with different scenarios and touch points of the customer journey in a commercial e-business platform for cosmetics products. It could be demonstrated that emotion recognition combined with eye- and gaze-tracking can be a powerful method for assessing situations and finding possible adaptations at application runtime. By the lab-based observation of users through multiple visual and physical channels we could establish a basis for improving the customer experience in the pre-purchase and after purchase phases, because the gathered knowledge can be used for optimizing a priori and persona-based adaptations and can lead to improved situation rules. A follow-up study that is currently under way will evaluate the effectiveness of the persona-based a priori and situation-aware runtime adaptations for the perceived individual customer experience.

In the future we will combine SitAdapt 2.0 functionality with web and big-data analytics to further improve the customer experience of e-business applications and to evaluate, which of the applied user monitoring technologies can be helpful in situation-aware end-user environments.

References

[1] Bissinger, B.C.: Messung und Analyse von bio-physischen und visuellen Daten zur Optimierung der User Experience und des Digitalmarketings, Master Thesis, Business Information Systems, Augsburg University of Applied Sciences, 2018

[2] Calvary, G., Coutaz, J., Bouillon, L. et al., 2002. “The CAMELEON Reference Framework”. Retrieved August 25, 2016 from http://giove.isti.cnr.it/projects/cameleon/pdf/CAMELEON%20D1.1RefFramework.pdf

[3] Chang, C.K.: Situation Analytics: A Foundation for a New Software Engineering Paradigm, IEEE Computer, Jan. 2016, pp. 24-33

[4] P. Ekman, “An argument for basic emotions,” Cogn. Emot., vol. 6, no. 3–4, pp. 169–200, 1992.

[5] Engel, J., Märtin, C., Forbrig, P.: Practical Aspects of Pattern-supported Model-driven User Interface Generation. Proc. HCI International 2017, Vancouver, Canada, 9-14 July, Vol. I, Springer LNCS, 2017, pp. 397-414

[6] Engel, J., Märtin, C., Forbrig, P.: A Concerted Model-driven and Pattern-based Framework for Developing User Interfaces of Interactive Ubiquitous Applications, Proc. First Int. Workshop on Large-scale and Model-based Interactive Systems, Duisburg, (2015), pp. 35-41

[7] Flach, J.M., Mulder, M, Van Paassen, M.M.: The Concept of the Situation in Psychology, in: Banbury, S. and Tremblay, S. (eds): A Cognitive Approach to Situation Awareness: Theory and Applications, Ashgate Publisching, Oxon (UK), (2004), pp. 42-60

[8] Galindo, J. et al.: Using user emotions to trigger UI adaptation, Proc. 12th Int. Conf. on Research Challenges in Information Science (RCIS), (2018)

[9] Halvorsrud, R., Kvale, K., Følstad, A.: Improving Service Quality through Customer Journey Analysis, J. of service theory and practice, vol. 26, 6, (2016), pp. 840-867

[10] Herdin, C., Märtin, C., Forbrig, P.: SitAdapt: An architecture for situation-aware runtime adaptation of interactive systems. Proc. HCI International 2017, Vancouver, Canada, 9-14 July, Vol. I, Springer LNCS, 2017, pp. 447-455

[11] Hibbeln, Martin; Jenkins, Jeffrey L.; Schneider, Christoph; Valacich, Joseph S.; and Weinmann, Markus. 2017. "How Is Your User Feeling? Inferring Emotion Through Human Computer Interaction Devices," MIS Quarterly, (41: 1) pp.1-21.

[12] Hudlicka, E. and M. D. Mcneese, “Assessment of user affective and belief states for interface adaptation: Application to an Air Force pilot task,” User Model. User-Adapt. Interact., vol. 12, no. 1, pp. 1–47, 2002.

[13] Lemon, K.N., Verhoef, P.C.: Understanding Customer Experience Throughout the Customer Journey, J. of Marketing: AMA/MSI Special Issue, Vol. 80 (Nov. 2016), pp. 69-97

[14] Märtin, C., Herdin, C.: Enabling Decision-Making for Situation-Aware Adaptations of Interactive Systems, Proc. 10th Forum Media Technology, FMT 2017, 29-30 Nov., St. Pölten, Austria, (2017)

[15] Märtin, C., Herdin, C., Engel, J.: Model-based User-Interface Adaptation by Exploiting Situations, Emotions and Software Patterns, Proc. CHIRA 2017, Funchal, Madeira, Portugal, 31 October-2 November, SCITEPRESS (2017), pp. 50-59

[16] Märtin, C., Rashid, S., Herdin, C.: Designing Responsive Interactive Applications by Emotion-Tracking and Pattern-based Dynamic User Interface Adaptation. Proc. of HCII 2016, Toronto, 17-22 July, Vol. III, Springer LNCS, 2016, pp. 28-36

[17] Meixner, G., Calvary, G., Coutaz, J.: Introduction to model-based user interfaces. W3C Working Group Note 07 January 2014. http://www.w3.org/TR/mbui-intero/. Accessed 27 May 2015

[18] Melchior, J., Vanderdonckt, J., Van Roy, P.: A Model-Based Approach for Distributed User Interfaces, Proc. EICS ‘2011, ACM (2011), pp. 11-20

[19] Meudt, S. et al.: Going Further in Affective Computing: How Emotion Recognition Can Improve Adaptive User Interaction, in: A. Esposito and L.C. Jain (eds.), Toward Robotic Socially Believable Behaving Systems – Vol. 1, Intelligent Systems Reference Library 105, Springer Int. Publishing Switzerland (2016), pp. 73-103

[20] Nasoz, F.: “Adaptive intelligent user interfaces with emotion recognition, University of Central Florida Orlando, Florida, 2004.

[21] Picard, R.:”Recognizing Stress, Engagement, and Positive Emotion”, Proc. IUI 2015, March 29-April 1, 2015, Atlanta, GA, USA, pp. 3-4

[22] Schilit, B.N., Theimer, M.M.: Disseminating Active Map Information to Mobile Hosts, IEEE Network, vol. 8, no. 5, pp. 22–32, (1994)

[23] Schmidt, A. Biosignals in Human-Computer Interaction, Interactions Jan-Feb 2016, (2016), pp. 76-79

[24] Yigitbas, E., Sauer, S., Engels, G.: A Model-Based Framework for Multi-Adaptive Migratory User Interfaces. In: Proceedings of the HCI 2015, Part II, LNCS 9170, Springer (2015), pp. 563-572

All Around Audio Symposium

List of Talks

Creative Strategies for Drawing and Animated Drawing in VR
Tania De León Yong, Academy of Media Arts Cologne & Eduardo Ortiz Vera, FAD Universidad Nacional Autónoma de México

Reality? Boring! Animation as an Audio-visual Catalyst of the Experienced World
Jens Meinrenken, Humboldt University Berlin

A Slice of Time Back into the Timeline: Some Considerations on Visual Rhythm in Animated Photography.
Rosangela De Araujo, Film University Babelsberg Konrad Wolf

Holistic Perception Through the Synthesis of Colour and Music
Victoria Wolfersberger, University of Applied Sciences Upper Austria, Hagenberg

Animation as Applied Art in the Field of Music Promotion, Analysis and Education
Iby-Jolande Varga, Independent Scholar

Evolving Opera – Immersive Technologies in Artistic Practice and Listening Experience
Maria Kallionpää, Hong Kong Baptist University & Hans-Peter Gasselseder, Aalborg University

Orchestrating Space by Icosahedral Loudspeaker (OSIL)
Gerriet K. Sharma & Frank Schulz, University of Music and Performing Arts Graz

Sound Quality and 3D-AUDIO
Friedrich Blutner, Synotec Psychoinformatik GmbH

The Secret to Great Video is Audio
Phillip Sonnleitner, MIKME

Rotting Sounds — Embracing the Temporal Deterioration of Digital Audio
Till Bovermann, University of applied arts Vienna & Almut Schilling, Academy of Fine Arts Vienna

On the Usage of Immersive Environmental Sound in VR for Clinical Purposes
Cornelius Pöpel, Ansbach University of Applied Sciences

Binaural Audio as Body Engineering
Martin Rumori, University of Music and Performing Arts Graz

The Acoustics Research Institute: Science Around Audio
Piotr Majdak, Acoustics Research Institute (ARI) of the Austrian Academy of Sciences

Diffuse Directivity
Paul Modler, Karlsruhe University of Arts and Design

Design and Implementation of the Laboratory for Immersive and Drone Based Journalism
Philipp Kessling, Hamburg University of Applied Sciences

t.b.a.
Thomas Görne, Hamburg University of Applied Sciences