Proceedings of the 7th Workshop on Interacting with Smart Objects

Florian Müller, Dirk Schnelle-Walka, Sebastian Günther, Karola Marky, Markus Funk, Max Mühlhäuser

Co-located with the *11th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS '19)*

Valencia, Spain

June 2019

Table of Contents

Preface

These are the proceedings of the *7th Workshop on Interacting with Smart Objects* (SmartObjects '19) in conjunction with EICS'19 held on June 18, 2019, in Valencia, Spain. This volume contains the ten accepted papers. Each submission was reviewed by three program committee members.

Objects that we use in our everyday life are ever-expanding their interaction capabilities and provide functionalities that go far beyond their original functionality. They feature computing capabilities and are, thus, able to capture, process and store information and interact with their environments, turning them into smart objects. Their wide range was covered by the submissions to this workshop. Smart objects know something about their users and, thus, allow for natural interaction. Natural interaction, in contrast, does not imply smartness. Smartness requires interaction with users and provides help. There are already commercialized products available that expose their properties and interaction capabilities. To enrich their potential and to lower affordances, they need to communicate to each other. Making sense out of the available data in this field is still an open research question. The overall goal should be to build an interactive ecosystem that (i) seamlessly discovers, connects and talks to its environment, (ii) is ubiquitous and (iii) allows the user to be in control.

The workshop examined these issues with regards to the following aspects:

- Smart Devices
- Smart Spaces

Putting together SmartObjects '19 was a team effort. We would like to send out our thanks to everybody who has helped us to organize this event:

- The authors, who have written and submitted their papers to the workshop.
- The program committee and the external reviewers, for their time and effort to write substantial and constructive review reports.

We hope that you will find this program interesting and thought-provoking and that the workshop will provide you with a valuable opportunity to share ideas with other researchers and practitioners from institutions around the world.

June 2019 Florian Müller
Darmstadt

Program Committee

Bo Begole	AMD, CA, USA
Marco Blumendorf	smartB, Germany
Oliver Brdiczka	Stella.ai, CA, USA
Jingyuan Chen	University of Science and Technology of China
Aba-Sah Dadzie	The Open University, United Kingdom
Boris Deruyter	Philips Research, Netherlands
Niloofar Dezfuli	TU Darmstadt, Germany
Markus Funk	TU Darmstadt, Germany
Tobias Grosse-Puppendahl	Porsche, Germany
Sebastian Günther	TU Darmstadt, Germany
Fahim Kawsar	TU Delft
Alexander Kröner	TH Nuernberg, Germany
Kris Luyten	Hasselt University, Belgium
Karola Marky	TU Darmstadt, Germany
German Montoro	Universidad Autónoma de Madrid, Spain
Max Mühlhäuser	TU Darmstadt, Germany
Florian Müller	TU Darmstadt, Germany
Patrick Reignier	Grenoble INP, France
Dirk Schnelle-Walka	modality.ai, CA, USA
Geert Vanderhulst	Alcatel-Lucent Bell Laboratories, Belgium
Alexandra Voit	University of Stuttgart
Raphael Wimmer	University of Regensburg, Germany
Massimo Zancanaro	Fondazione Bruno Kessler, Italy

External Reviewers

Kosch, Thomas	LMU Munich, Germany
von Wilmsdorff, Julian	Fraunhofer IGD, Germany

Author Index

Keyword Index

Mastering Musical Instruments through Technology in Solo Learning Sessions

Karola Marky
TU Darmstadt
Darmstadt, Germany
marky@tk.tu-darmstadt.de

Julien Gedeon
TU Darmstadt
Darmstadt, Germany
gedeon@tk.tu-darmstadt.de

Andreas Weiß
Musikschule Schallkultur
Kaiserslautern, Germany
andreas.weiss@musikschule-schallkultur.de

Sebastian Günther
TU Darmstadt
Darmstadt, Germany
guenther@tk.tu-darmstadt.de

Abstract
Mastering a musical instrument requires time-consuming practice even if students are guided by an expert. In the overwhelming majority of the time, the students practice by themselves and traditional teaching materials, such as videos or textbooks, lack interaction and guidance possibilities. Adequate feedback, however, is highly important to prevent the acquirement of wrong motions and to avoid potential health problems. In this paper, we envision musical instruments as smart objects to enhance solo learning sessions. We give an overview of existing approaches and setups and discuss them. Finally, we conclude with recommendations for designing smart and augmented musical instruments for learning purposes.

Author Keywords
Internet of Things; Assistance System; Musical Instruments, Learning Interfaces

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction
Learning a musical instrument is a time-consuming task that requires a lot of practice. Even if students are guided by an experienced musician or expert, the students prac-

tice by themselves in the overwhelming majority of the time since experienced musicians or experts are limited resources for several reasons. Those include availability and the cost of lessons. Traditional self-teaching materials, such as videos or textbooks, lack interaction and guidance possibilities. Thus, during solo practice sessions, the students do not receive immediate feedback. Adequate feedback, however, is highly important because the students might acquire wrong movements or postures that are difficult and time-consuming to correct later on. Missing feedback might even lead to health problems such as repetitive strain injury [18] or sore arms. Hence, students – especially beginners – require assistance during solo learning sessions that can provide them with adequate and immediate feedback.

More and more daily devices are turned into smart objects by either equipping them with sensors, actuators, and further computation and connection capabilities [13]. This technology can be leveraged to provide students with adequate feedback during solo learning sessions [9, 12] to enhance their music playing abilities. In this paper, we give an overview of existing technologies that augment and transform musical instruments into smart objects and derive recommendations for an optimal learning experience and accessibility during solo practice sessions.

Augmented and Smart Musical Instruments

In general, there are two possibilities to turn a musical instrument into a smart object: (1) augmentation by external devices and (2) integration of sensors and actuators. If the musical instrument is augmented by an external device, a conventional musical instrument, such as a guitar or a piano, it remains unaltered and the assistance comes from external devices. This could, for instance, be a screen [14], a projector [12, 15] or actuators [9]. If the musical instrument includes sensors and actuators, either a conventional musical instrument has to be transformed, or the musical instrument has to be built specifically to integrate the sensors and actuators. In the following, we give an overview of existing setups and discuss their benefits and drawbacks.

Screen-Based Augmentation

The simplest setup for screen-based augmentation is adding a screen that displays information. *Synthesia* [21] is a setup in which a screen is placed on top of a piano. The screen displays bars that are moving from the screen top towards the keys of a virtual piano. The length of the bar depicts the length of a tone, and the colors of the bars correspond to the hands. Numbers from 1 to 5 depict the finger that should press the respective key. While this visualization is widely adopted among autodidacts, a cognitive mapping from the keys on the screen to the real keys and fingers is required. This constitutes a difficulty because the student has to translate the depicted finger position that does not match his or her view. Information on the correct finger and body posture is not provided. The visual representation of tones as bars is more intuitive than sheet notation and can, therefore, be used more quickly [17].

Cakmakci *et al.* [2] use optical markers to detect finger positions on a bass guitar and overlay a camera image with target finger positions as well as directions. Liarokapis [11] introduces markers that can be placed in the student's environment. If the camera recognizes a marker, it augments it with finger placing information, e.g., a schematic depiction of a guitar chord. Thus, the student is not limited to look at one specific spot. The position tracking in these setups is limited to the position of the fingertips, but feedback regarding the finger posture, such as the angle of the joints or the required pressure, is not provided. Motogawa *et al.* [14] extended the setups mentioned above by adding a 3D-model of the ideal hand posture to the image. This introduces feedback regarding the finger posture and the students can check if their hand postures correspond to the presented one.

All setups presented up to now support the students in placing their fingers by showing them finger targets. Kerdvibulvech and Saito [10] display the currently played chord as well as target finger positions on a screen. Therefore, they track the guitar neck and the fingers of the students with markers. Because the students have to look at a screen, their view of the finger targets is inverted. Thus, similarly to traditional depictions on paper, a constant perceptual mapping between the screen and the real world is required.

Light-Based Augmentation
In light-based augmentation, visual light cues are used. There are three possible light sources: (1) an integrated light source, (2) a mounted light source and (3) projected light. Integrated light sources can, for instance, be built into the fretboard of a (bass) guitar [8, 20] or the key of a piano or keyboard [7]. A plethora of commercially available light-based products is already on the market, such as the *gTar* [8] or *FretLight* [20]. Mounted light sources can be mounted on conventional musical instruments [3]. The finger targets are visible and displayed at a position where no perceptual mapping is required.

In projection-based augmentation, the information is projected either directly on the surface of the instrument, or a projection surface close to the instrument. Takegawa and Terada [22] use a projector to support the learning of sheet music notation by projecting visual connections between written notes and piano keys.

The setup *guitAR* [12] projects finger targets on a guitar fretboard to show the students where to press. In the approach by Yang and Essl [24], *a Game of Tones* [15] and *P.I.A.N.O.* [17] a projection surface above a piano is used to project the next tones above the respective key. Both setups can be seen as an extension of *Synthesia* [21] to a projection surface. The projection surface, however, is not part of the piano and needs to be placed on top of it which restricts this technology to certain piano types. Furthermore, feedback regarding the finger and hand postures are not given in all setups.

In general, light-based setups cannot provide feedback regarding postures, since the students only know where to press but not how. The number of visible finger positions at a given time is limited because the student might not realize which key or string to press. The setups cannot react to the student's individual speed, since they can only display a sequence of chords or tones in a pre-defined tempo without adapting to the student's playing tempo.

Actuation- and Sensor-Based Augmentation
Augmentation by actuation means that a component sets the students' bodies into motion. The students' postures or movements can be captured by sensors. Similarly to light-based augmentation, the actuators and sensors can either be integrated into the musical instruments or can be mounted on them.

MusicJacket is a jacket that aims to teach body posture and the bowing techniques to violin players via vibrotactile feedback [23]. The students' motions are captured by several sensors while their movements are corrected by seven vibration motors on the upper body.

Mobile Music Touch (MTT) is a haptic glove that aims to teach the fingering for playing the piano [6]. While the student listens to a song, a vibration indicates the finger that is used to press this note.

The system *EMGuitar* uses electromyography to assess the student's finger postures on the guitar [9]. Electromyography captures the electrical activity that is produced by a muscle. Therefore, the student wears an armband-like construction that captures the muscle activity. Based on that it adjusts the tempo of displayed chords such that the student can play without interruptions.

Shin *et al.* propose a guitar setup that combines integrated LEDs, piezo sensors and microcontrollers [19]. They use different light colors for the different fingers and an application to provide feedback. The feedback is limited to the duration of a chord or tone and the position of the finger on the fretboard.

Recommendations

From the related work presented above and an expert discussion, we derive recommendations for designing and building smart musical instruments that are based on strings or keyboards:

1. Place guidance cues in the student's natural field of view.
2. Reduce or avoid perceptual mappings by placing the (visual) cues as close to the targets as possible, ideally directly on them. Furthermore, avoid inverted images.
3. Include postures, e.g. of the fingers or the body, in the cues.
4. Avoid additional (mounted) components, such as markers, because they might disturb or distract the students.
5. Wearable devices should not interfere with playing the musical instrument.
6. Be reactive on demand to adapt to the students' needs and the current learning curve (such as tempo).
7. Projection-based guidance and feedback should not require an additional (mounted) projection surface.

Discussion

New students are confronted with a rather huge cognitive load: they have to learn the setup of the musical instrument, instrument-specific music notation, postures of different body parts, and they have to develop a feeling for rhythm and tones. From a student's perspective, it makes sense to reduce this cognitive load as much as possible and gradually introduce one aspect after another instead of all at once. Smart and augmented musical instruments aim to support students by reducing the cognitive load in several ways:

- alternative and more intuitive visualizations of sheet music (cf. [21, 17])
- visual or haptic cues for finger and body placement (cf. [17])
- feedback regarding (in-)correct finger placement and body posture (cf. [23])

While those techniques are beneficial in the learning scenario, their long-term effects are not well understood yet. It remains unclear whether students might develop a dependency on those technologies, meaning they might overly rely on them. The purpose of the learning setup must be communicated clearly to the students which might result in overhead since using the setup requires introduction time and further practice.

The presented setups support students in different tasks but are mainly focused on finger positioning, i.e., they teach where to press a string or key but not how to do it. In this scope, direct approaches that do not require perceptual mappings are beneficial. The visual cues are at the real finger target and the students directly recognize whether they touch the correct spots. Wearing a head-mounted display that overlays the real world with the augmented information could be an appropriate solution for the target display. Since optical markers on the instrument and the fingertips might influence the student's movements, the finger position could be estimated by approaches that are based on depth-images [16].

Even if the students touch the correct spots, the finger, arm or body posture might be incorrect. Joints, especially those of the fingers, could be overstretched and the motions might be too cramped. Over a long period this might lead to persisting health problems and might even slow down further progress (in more complicated compositions).

Up to this point, we presented the means to enhance the learning of traditional musical instruments. But the advance of technologies enables a new generation of musical instruments. Musical instruments have already made several progressions such as the one from the harpsichord to the piano or from the baroque cello to the modern version of the cello. Screens, lights, and actuators bear the potential to transform existing musical instruments into new instruments with new features. For instance, the hand position problems that are present in nearly all musical instruments could be mitigated by sensor technology. This could, for instance, result in a piano with an ergonomic keyboard. Furthermore, learning-disabled musicians might benefit from simplified smart instrument setups as shown by Harrison *et al.* [5].

While our recommendations focus on solo learning sessions, augmented and smart musical instruments also enable new forms of remote collaboration. While approaches for remote music performances are already available in the literature [4, 1], remote music lessons nowadays rely on simple setups, such as video calling technology, that inherit several drawbacks from video learning materials although an experienced musician or expert is present and available. Sensors, lights, and actuators that are controlled by the music teacher might offer an enhanced learning experience. We envision that the teachers could, for instance, show a combination of tones on their musical instruments or play together with the students. The body and or finger postures of the experts are captured and presented to the students by the smart musical instrument in a way that requires minimal perceptual mapping. This would furthermore enable to take a lesson from a famous teacher that lives abroad.

Conclusion

In this paper, we gave an overview of existing setups for augmented and smart musical instruments. Those setups aim to support students in learning the respective instrument, especially beginners.

From the presented setups, we derive recommendations for designing and implementing augmented and smart musical instruments that optimize the learning experience. Furthermore, smart and augmented musical instruments can enable music lessons with teachers over distance with reduced requirements on perceptual mappings.

While this paper focuses on keyboard and string instruments, the extension to other types of musical instruments such as wind instruments constitutes an integral part of future work.

REFERENCES

1. Chrisoula Alexandrak and Rolf Bader. 2014. Using computer accompaniment to assist networked music performance. In *Proceedings of the Audio Engineering Society Conference (AES)*. Audio Engineering Society.
2. Ozan Cakmakci, François Bérard, and Joëlle Coutaz. 2003. An Augmented Reality based Learning Assistant for Electric Bass Guitar. In *Proceedings of the 10th International Conference on Human-Computer Interaction (HCI International)*.
3. LLC. Edge Tech Labs. 2019. FretZealot. http://fretzealot.com/. (2019). [Online; accessed: 14-March-2019].
4. Aristotelis Hadjakos, Axel Berndt, and Simon Waloschek. 2015. Synchronizing Spatially Distributed Musical Ensembles. In *Proceedings of the 12th Sound and Music Computing Conference (SMC '15)*.
5. Jacob Harrison, Alan Chamberlain, and Andrew P. McPherson. 2019. Accessible Instruments in the Wild: Engaging with a Community of Learning-Disabled Musicians. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19)*. ACM, New York, NY, USA, Article LBW0247, 6 pages. DOI: http://dx.doi.org/10.1145/3290607.3313037
6. Kevin Huang, Thad Starner, Ellen Do, Gil Weinberg, Daniel Kohlsdorf, Claas Ahlrichs, and Ruediger Leibrandt. 2010. Mobile Music Touch: Mobile Tactile Stimulation for Passive Learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 791–800.
7. CASIO America Inc. 2019. Casio LK-280. https://www.casio.com/products/electronic-musical-instruments/lighted-keys/lk-280. (2019). [Online; accessed: 14-March-2019].
8. Incident. 2019. gTar. www.incidentgtar.com. (2019). [Online; accessed: 14-March-2019].
9. Jakob Karolus, Hendrik Schuff, Thomas Kosch, Paweł W. Wozniak, and Albrecht Schmidt. 2018. EMGuitar: Assisting Guitar Playing with Electromyography. In *Proceedings of the Designing Interactive Systems Conference (DIS '18)*. ACM, New York, NY, USA, 651–655.
10. Chutisant Kerdvibulvech and Hideo Saito. 2007. Real-Time Guitar Chord Estimation by Stereo Cameras for Supporting Guitarists. In *Proceedings of the 10th International Workshop on Advanced Image Technology (IWAIT '07)*. 256–261.
11. Fotis Liarokapis. 2005. Augmented Reality Scenarios for Guitar Learning. In *Proceedings of the Third International Conference on Eurographics UK Theory and Practice of Computer Graphics*. Eurographics, 163–170.
12. Markus Löchtefeld, Sven Gehring, Ralf Jung, and Antonio Krüger. 2011. guitAR: Supporting Guitar Learning Through Mobile Projection. In *Proceedings of the Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 1447–1452.
13. David Molyneaux, Shahram Izadi, David Kim, Otmar Hilliges, Steve Hodges, Xiang Cao, Alex Butler, and Hans Gellersen. 2012. Interactive Environment-Aware Handheld Projectors for Pervasive Computing Spaces. In *Pervasive Computing*. Springer, 197–215. DOI: http://dx.doi.org/10.1007/978-3-642-31205-2_13

14. Yoichi Motokawa and Hideo Saito. 2006. Support System for Guitar Playing Using Augmented Reality Display. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '06)*. IEEE Computer Society, Washington, DC, USA, 243–244.

15. Linsey Raymaekers, Jo Vermeulen, Kris Luyten, and Karin Coninx. 2014. Game of Tones: Learning to Play Songs on a Piano Using Projected Instructions and Games. In *Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. ACM, New York, NY, USA, 411–414.

16. Jan Riemann, Martin Schmitz, Alexander Hendrich, and Max Mühlhäuser. 2018. FlowPut: Environment-Aware Interactivity for Tangible 3D Objects. *Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)* 2, 1, Article 31 (March 2018), 23 pages.

17. Katja Rogers, Amrei Röhlig, Matthias Weing, Jan Gugenheimer, Bastian Könings, Melina Klepsch, Florian Schaub, Enrico Rukzio, Tina Seufert, and Michael Weber. 2014. P.I.A.N.O.: Faster Piano Learning with Interactive Projection. In *Proceedings of the International Conference on Interactive Tabletops and Surfaces (ITS '14)*. ACM, New York, NY, USA, 149–158.

18. M. Rosety-Rodriguez, F. J. Ordóñez, J. Farias, M. Rosety, C. Carrasco, A. Ribelles, J. M. Rosety, and M. Gomez Del Valle. 2019. The Influence of the Active Range of Movement of Pianists' Wrists on Repetitive Strain Injury. *European Journal of Anatomy* 7, 2 (2019), 75–77.

19. Yejin Shin, Jemin Hwang, Jeonghyeok Park, and Soonuk Seol. 2018. Real-time Recognition of Guitar Performance Using Two Sensor Groups for Interactive Lesson. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '18)*. ACM, New York, NY, USA, 435–442. DOI: http://dx.doi.org/10.1145/3173225.3173235

20. The Fretlight Guitar Store. 2019. FretLight. https://fretlight.com/. (2019). [Online; accessed: 14-March-2019].

21. LLC. Synthesia. 2019. Synthesia – A fun way to learn how to play the piano. https://www.synthesiagame.com/. (2019). [Online; accessed: 14-March-2019].

22. Yoshinari Takegawa, Tsutomu Terada, and Masahiko Tsukamoto. 2012. A Piano Learning Support System Considering Rhythm. In *Proceedings of the International Computer Music Conference (ICMC)*. 325–332.

23. J. van der Linden, E. Schoonderwaldt, J. Bird, and R. Johnson. 2011. MusicJacket – Combining Motion Capture and Vibrotactile Feedback to Teach Violin Bowing. *IEEE Transactions on Instrumentation and Measurement* 60, 1 (Jan 2011), 104–113.

24. Qi Yang and Georg Essl. 2013. Visual Associations in Augmented Keyboard Performance. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME '13)*. 252–255.

Enabling Tangible Interaction through Detection and Augmentation of Everyday Objects

Thomas Kosch
LMU Munich
Munich, Germany
thomas.kosch@ifi.lmu.de

Albrecht Schmidt
LMU Munich
Munich, Germany
albrecht.schmidt@ifi.lmu.de

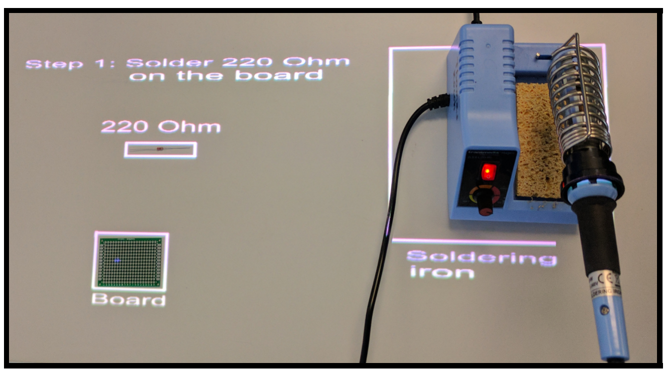


Figure 1: Augmented workspace in the context of electrical engineering. In-situ projections augment and trace objects as well as track user actions.

Abstract
Digital interaction with everyday objects has become popular since the proliferation of camera-based systems that detect and augment objects "just-in-time". Common systems use a vision-based approach to detect objects and display their functionalities to the user. Sensors, such as color and depth cameras, have become inexpensive and allow seamless environmental tracking in mobile as well as stationary settings. However, object detection in different contexts faces challenges as it highly depends on environmental parameters and the conditions of the object itself. In this work, we present three tracking algorithms which we have employed in past research projects to track and recognize objects. We show, how mobile and stationary augmented reality can be used to extend the functionalities of objects. We conclude, how common items can provide user-defined tangible interaction beyond their regular functionality.

Author Keywords
Object Tracking; Computer Vision; In-Situ Assistance; Assistive Systems; Workload-Aware Interfaces

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction

Augmenting common items with digitized content to extend their functionalities has been the focus of past research in the domain of tangible user interfaces [9]. Thereby, objects are tracked by a system that displays visual cues or extends the functionality of the object itself [10]. By rotating, repositioning, or placing objects in defined positions, user-defined actions can be triggered. Thus, common items are augmented by functionalities which they do not implement by themselves.

Two modalities to display such augmented content have emerged. Smart glasses, such as the Microsoft HoloLens[1], enable mobile use of augmented reality to display additional supporting content [4]. Furthermore, in-situ projection systems enable the augmentation of stationary workstations that can be used for practical exercises (see Figure 1). While smart glasses are preferred in a mobile context, in-situ projections are suitable for stationary settings. While mobile augmentation was preferred during practical physics exercises that required mobility of their students [18], industrial use cases [5] and social housing organizations [14, 15] found stationary settings more suitable. Furthermore, employing object augmentation provides cognitive alleviation, which has the potential to boost overall user performance and productivity [11, 12].

Both modalities use camera-based systems to recognize objects and enrich them with additional content. However, seamless object detection and augmentation poses challenges for different use cases. In this work, we present object detection strategies we employed in past research projects to enable object detection and augmentation. We discuss the advantages and disadvantages of different object tracking modalities. Finally, we present how user-defined tangibles from everyday items can be created by augmenting them with in-situ projections. We conclude with challenges that have to be considered when integrating ubiquitous object augmentation.

[1]www.microsoft.com/en-us/hololens - last access 2019-05-17

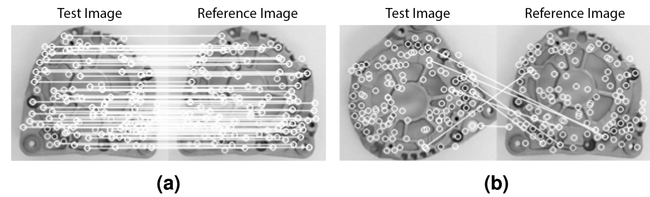


Figure 2: Object detection using SURF. The positioned object is compared to a reference image. Feature extraction, such as provided by the SURF algorithm, shows the similarity of the image. **(a):** Correct positioned image. **(b):** A rotated object does not guarantee that it will be detected relative to the reference image.

Object Tracking

To enable interaction with common items, suitable tracking systems and algorithms need to be employed. In the following, we present three object tracking strategies we have employed in past research.

SURF

The Speeded Up Robust Feature (SURF) algorithm [3] enables to recognize points and areas of "interest" in images. Due to its efficient implementation, it enables the processing of images in real-time. Thereby, the algorithm has been used for object detection by comparing points of interest in a captured image relative to a reference image [2]. SURF can be employed with inexpensive hardware since it processes color images. However, SURF is not rotation and

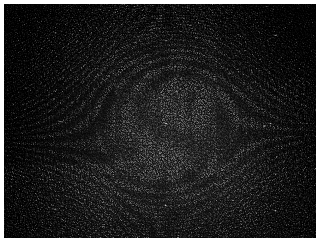


Figure 3: Infrared pattern of a Kinect v1 on a wooden plate [1].

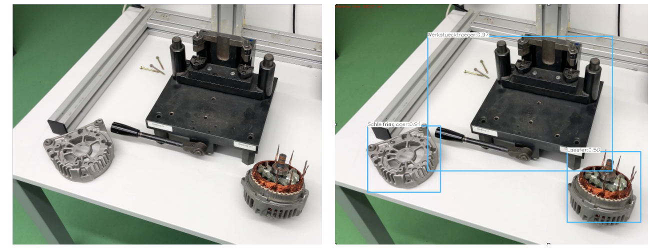
perspective invariant. This requires objects to be in a similar position that is expected by a system (see Figure 2).

Depth Sensing
A depth sensor, such as the Intel Realsense[2] or the Kinect v2[3], provide a 3D representation of objects that they are pointed to. Objects are recognized by analyzing the shape. Thereby, two relevant methods have emerged. The first method uses a projected infrared pattern on a surface (see Figure 3). Afterward, the depth sensor measures changes in the perspective of the pattern. This enables to detect the distance between infrared waves and allows a reconstruction of the 3D space on a surface [1]. The second method uses a Time-of-Flight approach. Thereby, the round trip time of an artificial light (i.e., infrared light) is measured between the sensor and a point on the surface. When the reflection of the light is captured, a 3D representation on of the surface is created [8].

Depth sensing is insensitive to lighting conditions. However, changes in perspective and rotation of objects may affect the overall detection quality. Thus, depth sensing is suitable for use cases where objects reside in stable positions.

You Only Look Once
The algorithm "You Only Look Once" (YOLO) is a deep learning approach to detect objects regardless of their perspective and position [17] (see Figure 4). It applies a single neural network on an image that detects features in bounding boxes after clustering their properties. By evaluating those properties, a probability of a correctly detected object is calculated. While YOLO represents a robust real-time method to detect objects regardless of their positioning and

[2] www.intel.com/content/www/us/en/architecture-and-technology/realsense-overview.html - last access 2019-05-17
[3] https://developer.microsoft.com/en-us/windows/kinect - last access 2019-05-17



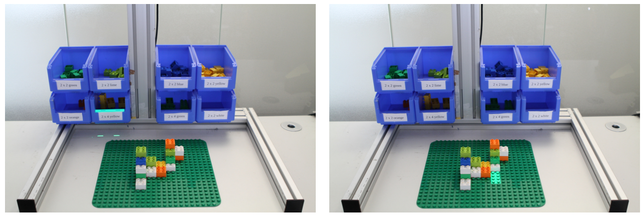
(a) (b)

Figure 4: Using YOLO to detect objects independent from their position. **(a):** Test image to evaluate a trained model. **(b):** Detected objects using YOLO. A blue bounding box denotes the detected objects.

perspective, it requires an extensive training set beforehand. Furthermore, training a neural network on a large data set requires time and, depending on the use case, fast computational hardware to speed up the training process.

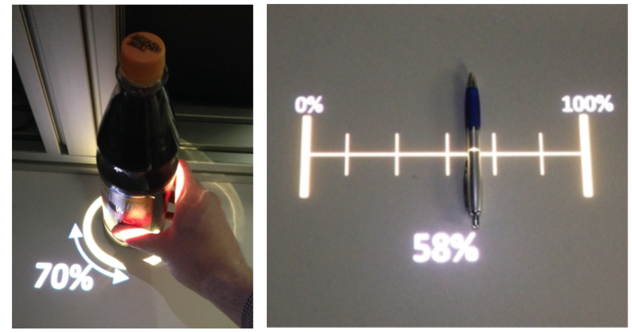
Object Augmentation
Objects can be used as a visual cue for interaction or interaction device itself. In the following, we show implementations of tangible object augmentation we have conducted in the past.

Ambient Augmentation
After recognizing the type of object, cues can be used to implicitly guide the user through a series of actions. Figure 5 shows an augmented workspace that uses in-situ projection as a guide through a series of assembly steps. By detecting the user's action and items on the workspace, in-situ projections are placed on the current relevant bin or final spot for assembly. While boosting the overall performance of workers in industrial environments [7], people with



(a) (b)

Figure 5: Augmenting a workplace using in-situ projections. **(a):** A detected item selection bin is visually highlighted. **(b):** A projection on the working area depicts the final position of an assembly part.



(a) (b)

Figure 6: User-defined tangibles that use in-situ projections to provide feedback. **(a):** Rotating a bottle similar to a knob. **(b):** Using a pen as a slider [6].

dementia and loss in memory benefit from in-situ projections [13].

User-defined Tangibles
Regular objects can be registered as user-defined tangible that is made available for interaction [6]. For example, rotating (see Figure 6a) or positioning (see Figure 6b) objects can be used to change the speaker volume.

After registering the object, a series of options are made available to the user. The user can choose to interact with existing objects or register new objects. Such objects can be everyday items which do not implement a logic. This transforms objects into user-defined tangibles that are already around the user with just-in-time interaction.

Challenges and Future Work

Seamless object detection and augmentation in home and workplace settings are prone to certain challenges. In this work, we presented three strategies to detect objects and augment objects. However, choosing the right detection modality depends on the environment as well as on the properties of the object itself. For example, a depth sensor will struggle to detect flat objects as they have scarce 3D properties. While a regular color camera can solve this problem, it is sensitive to the overall environmental illumination. In future work, we want to combine the definition and detection of user-defined tangibles by using an approach that combines color as well as depth images [16]. Thereby, a combination of depth and color data provides an approximation of object type.

Furthermore, privacy and ethical considerations have to be taken into account. By using the presented camera-based approach, public and private spaces are recorded during user interaction. While users can give consent to process the collected data in private settings, public spaces and workplaces are more sensitive to privacy-related issues. In future work, we want to investigate those ethical ramifications. Ultimately, we will investigate design guidelines that explore how a camera-based approach can be conducted while minimally invading the user's privacy.

Conclusion

In this work, we present three strategies to detect objects which we have employed in past research projects. We outline the advantages and disadvantages of each strategy which we have encountered. We show how object detection and user-defined tangibles can be implemented to provide ambient or explicit interaction. Finally, we discuss challenges that have to be tackled before enabling seamless object tracking in home and work settings. Since common objects do not implement any logic, we believe that external object augmentation paves the way for ubiquitous tangible interaction at home, public spaces, and workplaces.

REFERENCES

1. Michael Riis Andersen, Thomas Jensen, Pavel Lisouski, Anders Krogh Mortensen, Mikkel Kragh Hansen, Torben Gregersen, and Peter Ahrendt. 2012. Kinect depth sensor evaluation for computer vision applications. *Aarhus University* (2012), 1–37.
2. Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. 2006. SURF: Speeded Up Robust Features. In *Computer Vision – ECCV 2006*, Aleš Leonardis, Horst Bischof, and Axel Pinz (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 404–417.
3. Dimitris Bouris, Antonis Nikitakis, and Ioannis Papaefstathiou. 2010. Fast and Efficient FPGA-Based Feature Detection Employing the SURF Algorithm. In *2010 18th IEEE Annual International Symposium on Field-Programmable Custom Computing Machines*. 3–10. DOI:http://dx.doi.org/10.1109/FCCM.2010.11
4. Wilhelm Dangelmaier, Matthias Fischer, Jürgen Gausemeier, Michael Grafe, Carsten Matysczok, and Bengt Mueck. 2005. Virtual and augmented reality support for discrete manufacturing system simulation. *Computers in Industry* 56, 4 (2005), 371 – 383. DOI: http://dx.doi.org/https://doi.org/10.1016/j.compind.2005.01.007 The Digital Factory: An Instrument of the Present and the Future.
5. Markus Funk, Andreas Bächler, Liane Bächler, Thomas Kosch, Thomas Heidenreich, and Albrecht Schmidt. 2017. Working with Augmented Reality? A Long-Term Analysis of In-Situ Instructions at the Assembly Workplace. In *Proceedings of the 10th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, New York, NY, USA. DOI: http://dx.doi.org/10.1145/3056540.3056548
6. Markus Funk, Oliver Korn, and Albrecht Schmidt. 2014. An Augmented Workplace for Enabling User-Defined Tangibles. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA'14)*. ACM, New York, NY, USA, 1285–1290. DOI: http://dx.doi.org/10.1145/2559206.2581142
7. Markus Funk, Thomas Kosch, and Albrecht Schmidt. 2016. Interactive Worker Assistance: Comparing the Effects of In-situ Projection, Head-mounted Displays, Tablet, and Paper Instructions. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (2016), 934–939. DOI: http://dx.doi.org/10.1145/2971648.2971706
8. Burak Gokturk, Hakan Yalcin, and Cyrus Bamji. 2004. A Time-Of-Flight Depth Sensor - System Description, Issues and Solutions. In *2004 Conference on Computer Vision and Pattern Recognition Workshop*. 35–35. DOI:http://dx.doi.org/10.1109/CVPR.2004.291

9. Eva Hornecker and Jacob Buur. 2006. Getting a Grip on Tangible Interaction: A Framework on Physical Space and Social Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. ACM, New York, NY, USA, 437–446. DOI: http://dx.doi.org/10.1145/1124772.1124838

10. Martin Kaltenbrunner and Ross Bencina. 2007. reacTIVision: A Computer-vision Framework for Table-based Tangible Interaction. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction (TEI '07)*. ACM, New York, NY, USA, 69–74. DOI: http://dx.doi.org/10.1145/1226969.1226983

11. Thomas Kosch, Yomna Abdelrahman, Markus Funk, and Albrecht Schmidt. 2017. One Size does not Fit All - Challenges of Providing Interactive Worker Assistance in Industrial Settings. *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (2017), 6. DOI: http://dx.doi.org/10.1145/3123024.3124395

12. Thomas Kosch, Markus Funk, Albrecht Schmidt, and Lewis Chuang. 2018. Identifying Cognitive Assistance with Mobile Electroencephalography: A Case Study with In-Situ Projections for Manual Assembly. In *Proceedings of the 10th ACM SIGCHI symposium on Engineering interactive computing systems*. ACM. DOI: http://dx.doi.org/10.1145/3229093

13. Thomas Kosch, Romina Kettner, Markus Funk, and Albrecht Schmidt. 2016. Comparing Tactile, Auditory, and Visual Assembly Error-Feedback for Workers with Cognitive Impairments. In *Proceedings of the 18th international ACM SIGACCESS conference on Computers & accessibility*. ACM. DOI: http://dx.doi.org/10.1145/2982142.2982157

14. Thomas Kosch, Kevin Wennrich, Daniel Topp, Marcel Muntzinger, and Albrecht Schmidt. 2019. The Digital Cooking Coach: Using Visual and Auditory In-Situ Instructions to Assist Cognitively Impaired during Cooking. In *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, New York, NY, USA. DOI:http://dx.doi.org/10.1145/3316782.3321524

15. Thomas Kosch, Pawel Wozniak, Erin Brady, and Albrecht Schmidt. 2018. Smart Kitchens for People with Cognitive Impairments: A Qualitative Study of Design Requirements. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA. DOI: http://dx.doi.org/10.1145/3173574.3173845

16. Kevin Lai, Liefeng Bo, Xiaofeng Ren, and Dieter Fox. 2011. Sparse distance learning for object recognition combining RGB and depth information. In *2011 IEEE International Conference on Robotics and Automation*. 4007–4013. DOI: http://dx.doi.org/10.1109/ICRA.2011.5980377

17. Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You Only Look Once: Unified, Real-Time Object Detection. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

18. Martin P. Strzys, Sebastian Kapp, Michael Thees, Pascal Klein, Paul Lukowicz, Pascal Knierim, Albrecht Schmidt, and Jochen Kuhn. 2018. Physics holo.lab learning experience: using smartglasses for augmented reality labwork to foster the concepts of heat conduction. *European Journal of Physics* 39, 3 (mar 2018), 035703. DOI: http://dx.doi.org/10.1088/1361-6404/aaa8fb

Improving Presence Detection For Smart Spaces

Julian von Wilmsdorff
Technische Universität Darmstadt
Darmstadt, 64289, Germany
julian.von.wilmsdorff@igd.fraunhofer.de

Biying Fu
Fraunhofer Institute for Computer Graphics Research IGD
Darmstadt, 64289, Germany
biying.fu@igd.fraunhofer.de

Florian Kirchbuchner
Fraunhofer Institute for Computer Graphics Research IGD
Darmstadt, 64289, Germany
florian.kirchbuchner@igd.fraunhofer.de

Abstract
In this paper, we present a novel sensor for smart spaces based on electric field sensing. It detects and classifies several events around a door to improve presence detection. We are able to detect events including *inside*, *outside*, *entry*, *exit* and *none*. In contrast to photoelectric sensors, it does not require a direct line of sight and also does not react to objects like suitcases with wheels or similar things like wheelchairs. Based on a conducted test study with 12 participants, we showed that we are able to detect the given classes with an overall accuracy of 90.3 %.

Author Keywords
Sensors; Smart Spaces; Electric Field Sensing; Signal Processing

ACM Classification Keywords
Hardware [Sensors and actuators]; Hardware [Digital signal processing]

Motivation
In modern smart spaces, the information of the presence of users is mandatory for many systems. By knowing the number of users in a room, smart objects can adapt their behaviour to fit the current situation. For example, lights can be turned off in case no persons are present to save energy or music speakers can increase the volume when

more persons enter the room.

Commonly used presence detectors that are based on infrared detection are not sufficient for this application. If a person enters a room and remains calm, the sensor has no means to know if the person left the room or is sitting nearly motionless in the room.

Optical barriers can improve this situation, but do not cover other aspects of real life situations. If two light barriers are placed at every entrance of a room, directional information of exit- and enter-events can be calculated. But optical systems lack the capability of differentiating between objects and persons.

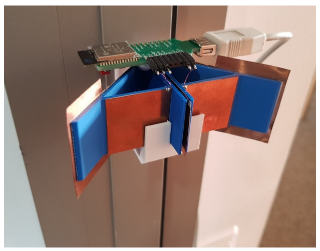


Figure 1: sensor and copper electrodes placed on door

This is why we implemented a directional sensor (shown in Figure 1) based on electric field sensing. These sensors react very sensitive to steps. That is the reason why objects with wheels are not recognized by the detection algorithm. In addition, compared to mere capacitive sensors, our passive electric field sensors have a higher range.

Related Work

The principle of electric field sensing is well known for over hundred years, but lots of application areas have been revived in the last few decades with emerging new processing algorithms and sensor designs. This technology gained lots of popularity in sense of low power consumption, no emission of electrical fields and high privacy preserving aspects. In the medical domain, applications like remote EEG measurement has been implemented by Prance et al. [6]. The group of Wilmsdorff et al. [7] have showed in their research paper lots of exploratory experiments for different use cases, for example no-touch gesture recognition for wearables and traffic observation using electrical field sensing. Xavier et al. [4] worked with the possibility of using this technology for indoor positioning and even person recognition based on gait patterns on two different days. Similar work for indoor positioning system using electrical potential sensing on a smart floor has been presented by Fu et al. [2]. Cohn et al. [1] made some efforts by applying this technology in gaming context. They augmented a customized gaming pad into a device with multiple input modalities like jumping and stepping without using the control stick on the gaming pad. Examples of wearables based on electric field sensing that can detect movements of legs and even the touch of human hair is shown by [5].

Door as an entry point to a secured location is quite interesting to interact with. Gjoreski et al. [3] showed in their work that it is possible to identify person by just analyzing door accelerations in Time and Frequency domain. In the following sections, we present a novel use-case of electrical potential sensing to be a smart presence detector. We first introduce the hardware implementation, followed by the detection algorithm and finally conclude our findings in the evaluation section.

Hardware Implementation

The sensor contains four core components. These components are:

- A UART to USB bridge for communication purposes
- An ESP32 micro controller of which two ADCs are used in 12bit mode
- Two electric field sensing groups
- Two shielded electrodes for every sensing group

A measurement group consists of an instrumentation amplifier, which meters the voltage between two pre-charged electrodes. To pre-charge the electrodes, half of the supply voltage is linked to both electrodes over two 1GΩ resistors. The current running through these resistors slowly pulls the measured signal back to a defined voltage level, removing some of the wanted signal in the process. To prevent a too strong loss of the signal, these resistors have to have a high value. Omitting these resistors would result in a higher range and increased sensitivity of the sensor, but would also introduce the problem of railing voltages. This happens if a voltage over the supply voltage (3.3V) or a negative voltage is created between the electrodes. Without pre-charging of the electrodes, the voltage level would not (or very slowly) recover to a range measurable by the ADC of the micro controller. By tying the potential of the electrodes to 1.65V, the sensor values will normalize within seconds, even if railing occurred. Figure 2 shows the simplified circuit of a measurement group. If the voltage of the first electrode is p_a and the voltage of the second electrode is p_b, the voltage u given by the instrumentation amplifier will be:

$$u = \frac{1}{2}V_{cc} + (p_a - p_b)$$

The voltage u is sampled by an ADC of the micro controller and further processed. This voltage is influenced by movements of the human body. Since there is a tiny amount of charge on the body, it will attract the opposite charge on the electrodes while approaching the sensor, but not the same amount on every electrode because of the arrangement of the electrodes. The induced potential difference between both electrodes is the input for our instrumentation amplifier.

Detection Algorithm

Since the sensor consists of two measurement modules, every module will output its own measurements. The measurement modules use a scan frequency of 50 Hz, the frequency of the European power grid. In this way, noise created by power outlets and power lines is suppressed by under-sampling.

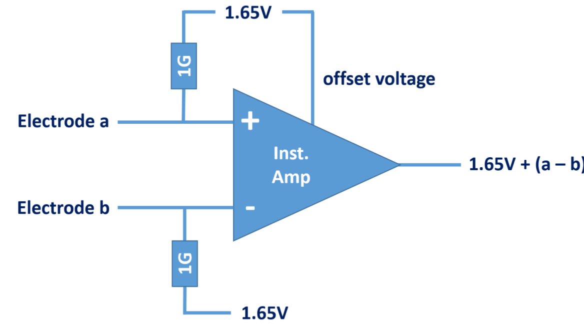


Figure 2: simplified circuit of a measurement group

The two outputs of the sensor will be processed by a pipeline. Every module uses a 12 bit ADC, which is equal to values from 0 (= 0V) to 4095 (= 3.3V). Because of the pre-charging of the electrodes, the normal baseline of a module is 2048, around half of the measurement range. Due to variances of the electrical components and environmental conditions like air humidity and temperature, the baseline can have an offset up to 10% of the original 2048. This is why the first stage of the pipeline is to calculate the real baseline of every measurement module and subtract it so that the values are zero based. This stage will only be active if there were no activities for at least 25ms. Otherwise the sensor would calibrate its baseline to the level of human steps.

The second stage is to form the first derivation of the two signals. This is needed to calculate the moment when the feet of a person hit the ground, which is represented by a local minimum or maximum. Note that no information can be obtained by the distinction of minima and maxima, because this only depends on the charge of a person. If a person is charged negatively, their steps will give a negative amplitude, otherwise a positive. The position of the extremum will be stored, but only if the following conditions are met:

- The first derivation is crossing the zero line. The direction of the crossing does not matter out of the stated reasons.
- The amplitude of the signal has to overcome a certain threshold. Simple noise will be discarded this way.
- The extremum has a certain minimal euclidean distance to the previous extremum. This way, if a single extremum that was corrupted by noise would appear as two or more extrema, the algorithm will only note one extremum.

This stage only operates on the previously calculated positions of the extrema. If no new peaks are detected for at least 25ms, the third stage of the pipeline is processed. For each peak we compute the sign of the difference in amplitude of the two signals. The electrodes of the sensor are placed in such a way that the position of a person in relation to the sensor will give a stronger signal in one measurement module, depending on if the person is moving on the right of the sensor or on the left. When calculated for each peak, this stage of the pipeline will result in a sequence of negative or positive peaks. In a best case scenario, a person which is moving from right to left would give the sequence: +1, +1, ..., +1, -1, -1, ..., -1. Note that the number of peaks is determined by the number of steps of the detected person.

The fourth and last stage of the algorithm is an auto correlation. Four different cases of sequences are evaluated:

- $\{+1, +1, ..., +1, -1, -1, ..., -1\}$: The person is moving from right to left
- $\{-1, -1, ..., -1, +1, +1, ..., +1\}$: The person is moving from left to right
- $\{+1, +1, ..., +1\}$: The person is moving only on the right side of the sensor
- $\{-1, -1, ..., -1\}$: The person is moving on the left side of the sensor

These are ideal sequences. A normal given sequence could contain outliers that obfuscate the sequence. To eliminate those, every +1 or -1 that has no adjacent peak with the same sign will be discarded. For example, the sequence $\{+1, +1, +1, -1, +1, -1, -1\}$ would result in $\{+1, +1, +1, -1, -1\}$. The auto correlation is only computed if three or more peaks are contained in the reduced sequence. Otherwise the result will be unreliable. Such weak signals are discarded because they originate most likely of persons moving at a large distance of the sensor or noise. In these cases, the algorithm will output the none-class. If there are enough peaks, the auto correlation matches the reduced sequence with these four functions:

- person moving left to right: modified Heaviside step function
$$H_1(x) = \begin{cases} -1 & x \leq 0 \\ 1 & x > 0 \end{cases}$$
- person moving right to left: inverted modified Heaviside step function
$$H_2(x) = \begin{cases} 1 & x \leq 0 \\ -1 & x > 0 \end{cases}$$
- person moving on the right: constant positive function
$$P(x) = 1$$
- person moving on the left: constant negative function
$$N(x) = -1$$

Entry	Straight line from Position 1 to Position 4
Exit	Straight line from Position 4 to Position 1
Outside	Starting from Position 1 to circle around Position 2 and return back to Position 1
Inside	Starting from Position 4 to circle around Position 3 and return back to Position 4

Table 1: The selection of pre-marked paths regarding the different classes has been given and each path was taken twice.

The function with the lowest error will be selected and represents the final result of the algorithm.

Evaluation

To illustrate the proof of concept, we conducted a test study with 12 participants. The participants have an average height of 174.9 cm ranging from 163 to 186 cm and contain 5 females and 7 males. We asked the participant to walk on predefined paths as shown in Figure 3. Each path were taken twice to determine the 5 different target classes of {*inside*,*outside*,*exit*,*entry*, *none*}. The approximate sensing range is indicated by the area of the blue circle. Four positions from 1 to 4 have been marked to indicate the path. The walking speed is not constrained. The walking direction was instructed as given in Table 1.

We noted the success- and mismatch-rate for each run to derive the confusion matrix shown in Figure 4.

We did an additional experiment to show that electric field sensing in contrast to photoelectric sensors will not be disturbed by objects. In Figure 5, we plotted two different signals. The upper plot shows the signal, when a person is entering the room rolling a wheel chair, while the plot below shows the signal when a person is entering the room without any objects. As shown, the signals are nearly identical

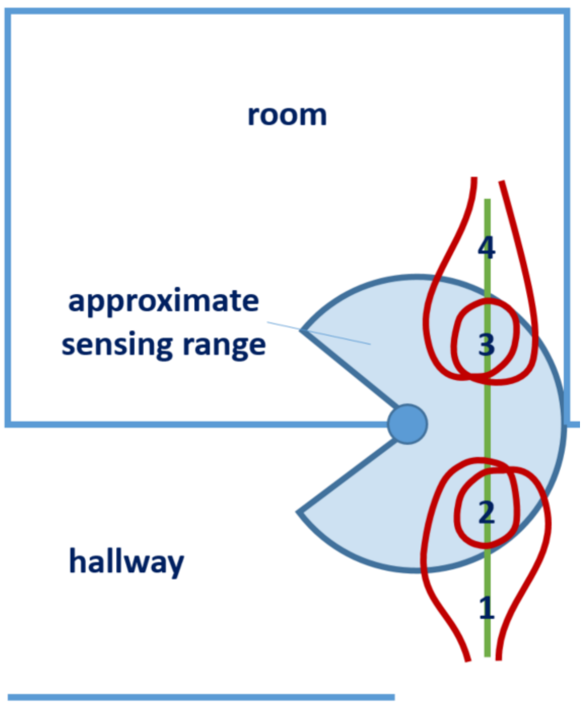


Figure 3: evaluation setup and walking paths

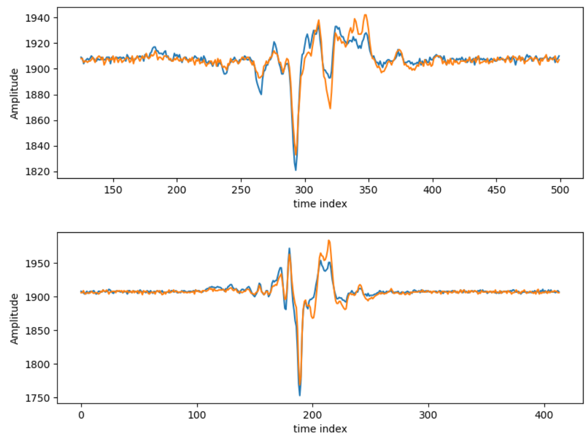


Figure 5: signal of entering the room with (above) and without (below) wheelchair

Figure 4: confusion matrix of the five different classes

and do not contain any features indicating another moving object. Both entry events were classified correctly by the sensor.

Conclusion & Future Work

We presented a novel approach for counting exit- and entry-events with a sensor based on electric field sensing. The evaluation shows that this concept is promising. To improve the performance even more, the placement of the sensor could be further examined and optimized. An important point would be to enhance the implementation to detect multi-user scenarios. Regarding the advantages of this technology like low power consumption, no need for direct line of sight and insensitivity to objects, this technology is very suitable for this use-case.

REFERENCES

1. Gabe Cohn, Sidhant Gupta, Tien-Jui Lee, Dan Morris, Joshua R. Smith, Matthew S. Reynolds, Desney S. Tan, and Shwetak N. Patel. 2012. An Ultra-low-power Human Body Motion Sensor Using Static Electric Field Sensing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. ACM, New York, NY, USA, 99–102. DOI: http://dx.doi.org/10.1145/2370216.2370233

2. Biying Fu, Florian Kirchbuchner, Julian von Wilmsdorff, Tobias Grosse-Puppendahl, Andreas Braun, and Arjan Kuijper. 2019. Performing indoor localization with electric potential sensing. *Journal of Ambient Intelligence and Humanized Computing* 10, 2 (01 Feb 2019), 731–746. DOI: http://dx.doi.org/10.1007/s12652-018-0879-z

3. Hristijan Gjoreski, Rok Piltaver, and Matjaz Gams. 2015. Person Identification by Analyzing Door Accelerations in Time and Frequency Domain. In *Ambient Intelligence - 12th European Conference, AmI 2015, Athens, Greece, November 11-13, 2015, Proceedings*. 60–76. DOI: http://dx.doi.org/10.1007/978-3-319-26005-1_5

4. Tobias Grosse-Puppendahl, Xavier Dellangnol, Christian Hatzfeld, Biying Fu, Mario Kupnik, Arjan Kuijper, Matthias Hastall, James Scott, and Marco Gruteser. 2016. Platypus - Indoor Localization and Identification through Sensing Electric Potential Changes in Human Bodies. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '16)*. ACM, New York, NY, USA, 14. DOI: http://dx.doi.org/10.1145/2906388.2906402

5. A Pouryazdan, RJ Prance, H Prance, and D Roggen. 2016. Wearable electric potential sensing: a new modality sensing hair touch and restless leg movement. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, 846–850. dx.doi.org/10.1145/2968219.2968286

6. R. J. Prance, S. T. Beardsmore-Rust, P. Watson, C. J. Harland, and H. Prance. 2008. Remote detection of human electrophysiological signals using electric potential sensors. *Applied Physics Letters* 93, 3 (2008).

7. Julian von Wilmsdorff, Florian Kirchbuchner, Biying Fu, Andreas Braun, and Arjan Kuijper. 2019. An experimental overview on electric field sensing. *Journal of Ambient Intelligence and Humanized Computing* 10, 2 (01 Feb 2019), 813–824. DOI: http://dx.doi.org/10.1007/s12652-018-0877-1

MODS: Modularly Operated Digital Signage

Rolf Egert
Technische Universität Darmstadt
Darmstadt, Germany
egert@tk.tu-darmstadt.de

Tim Grube
Technische Universität Darmstadt
Darmstadt, Germany
grube@tk.tu-darmstadt.de

Max Mühlhäuser
Technische Universität Darmstadt
Darmstadt, Germany
max@tk.tu-darmstadt.de

Abstract

Signage is transitioning from static analogue signs towards Digital Signage (DS), which introduces a variety of benefits. Among those are remote-maintenance, supporting dynamic content like videos and animations, and the simplification of updating content. However, DS solutions, despite being ubiquitous, are often tailored to specific use-cases, which limits their re-usability and updateability in case of severe changes to their environment. For instance, digital door signs for office spaces may become unusable if the office space is reorganized as storage or seminar rooms. Coping with such changes may result in additional costs since new DS solutions need to be purchased or in-depth changes to the software of the currently deployed DS solution are required. To address these problems we propose the Modularly Operated Digital Signage (MODS) framework, facilitating the dynamic changing of DS functionality in a modular fashion. We present the frameworks modular concept and describe its individual components. Subsequently, we briefly elaborate on the properties of the currently implemented modules. Additionally, we discuss the conducted pre-study to receive a first indicator for the usability of the framework.

Author Keywords

Digital Signage; User Interaction; Modularity

ACM Classification Keywords

H.4.m [Information systems applications:]: Miscellaneous

Introduction

Using signs for displaying, highlighting and conveying information is an established technique that is ubiquitous in our daily life [2]. These signs can range from small warning symbols, which are printed on gadgets and products, over door-signs used for identifying rooms in buildings, towards large scale advertisements. While some of those signs serve their purpose for a long time period (e.g., organization names, street signs), others may have shorter life spans, resulting in outdated or false displayed information. The consequences of the resulting misinformation can range from minor irritation to misunderstandings, which ultimately can be harmful if signs that are supposed to inform about dangerous aspects (e.g., dangerous locations) are not up-to-date. To address the problem of outdated information, signs need to be replaced regularly, which involves printing and the manual installation of the updated sign at all positions where it should be displayed. Nowadays, digitization aims at simplifying the task of displaying and updating information by using Digital Signage (DS) instead of static analogue ones. Alongside this transition towards DS come additional benefits. Among those are the capability of displaying dynamic content like animations or videos, the integration of sound, remote updateability and many more [3].

However, many solutions are still tailored towards a specific use-case, like showing a playlist of pictures and videos or presenting a web-page. One example of such an approach is Xogo [5], which allows displaying a predefined set of images and movies. However, the content is rather static and the user interaction is limited to change between the content locally available on the device. Other solutions that aim to increase the flexibility of interaction and functionality are based on Raspberry Pis. Among those solution are *piSignage* [1], which also facilitates playlist-based functionalities but allows for displaying additional content like web-pages or calendar events. Another approach is presented by the authors of [6], where Raspberry Pis are used to support the room management in university buildings and provide auxiliary information relevant for students. Other solutions are more powerful but are only available as commercial products. An example of such a system is [4]. The tool provides sophisticated functionalities, supporting remote-control via App, providing touch interaction for users and also enables querying information of different formats. The whole suit is controlled using a web interface. However, the tool is closed-source and, therefore, its extensibility is limited.

One of the main issues identified from related solutions is the limitation that DS is often tailored towards specific use-cases. This still limits the usability of many DS applications to the very domain they were developed for, which is problematic in case of severe changes happen to the environment. For instance, an office environment is rearranged as a combination of storage and lecture rooms, or additional use-cases were identified that are relevant for the current environment and are not covered by the functionality of the currently deployed software. If the previously installed DS does not support adequate functionality to fulfill the requirements, they need to be replaced or more drastic modifications have to be made to the software that is running on those devices. Consequently, this may lead to significant monetary investments, either for buying new products or for modifying the software and deploying it on all available devices.

In this context, we propose the Modularly Operated Digital Signage (MODS) framework, which addresses the missing dynamic adaption to multiple use-cases by organizing different functionality in a plugin-based fashion. This allows to easily change the appearance and functionality of the DS according to the current use case. Moreover, MODS introduces the concept of Interrupt Handlers as an event-based interaction mechanism independent of the current use-case. Additionally, we evaluated the usability of MODS in a pre-study.

The remaining paper is structured as follows. First, the general concept of MODS is explained and the currently implemented modules are briefly introduced. Second, the conducted pre-study assessing the usability of MODS is explained. Finally, the work is concluded and an outlook to future work is given.

MODS Concept

The purpose of MODS is twofold: first, it aims to provide the functionality that is generally supported by DS, like displaying dynamic content, enable user interaction and updating information; second, the framework addresses the problem of dynamically adapting to changing use-cases of the DS by enabling the exchange of functionality. Note that hereby two types of users have to be distinguished. One user is interacting with the DS to receive information and the second one (i.e., administrative user) is using the framework to display and update information. To facilitate the adaption of the DS towards multiple use-cases, the framework allows changing the running software to fit the desired use-case by loading the corresponding functionalities as modules. The general structure of the MODS framework is displayed as Fig. 1

The framework is separated into two main parts, namely *Sign Program*, which describes the software part that is running on the DS device and the *Update Program*, which runs on the users/administrators computer enabling remote

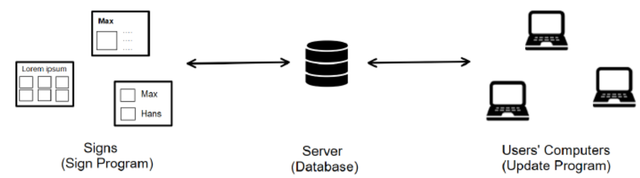


Figure 1: General three component concept of MODS

configuration and control. In the following, these two parts are introduced in more detail.

Sign Program

The *Sign Program* represents the core of the MODS framework and acts as the program that is running on the device used as a digital sign. This part of the framework allows realizing the end-user application (i.e., the visual and interactive part of DS which is perceived by the person using it) of the DS for a diverse set of use-cases in a plugin-based fashion. The available plugin types and their connections are described in the following in more detail.

Mode

A *mode* defines a set of use-cases the DS addresses and, therefore, provides a set of functionalities that are related to these use-cases. At each point in time, only one mode can be active on a DS. An exemplary mode can be a “booking mode” for seminar rooms that allows reserving dedicated time slots for using the room. While this mode is running, another mode cannot be executed simultaneously. In this scenario, a use-case represents the general application scenario where the DS is deployed, like an office situation or as an advertisement board. Each mode can be loaded in a plugin-based fashion, facilitating a simple way of changing the set of functionalities which is required by the DS. The visualized components of a mode are defined as so-

Figure 2: Detailed information view for individual people listed in the Office Mode.

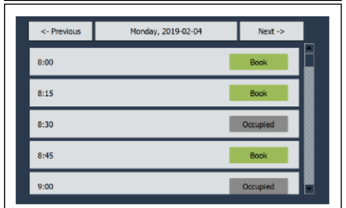


Figure 3: Conference View which enables the booking of individual time-slots at specific dates.

called *SignViewComponents*, which are part of different *SignViews* of a mode and are responsible for displaying arbitrary content. The SignViewComponents can receive interaction information from the user of the DS and forward it to the internal logic for further processing.

SignViews & SignViewComponents

The SignViews represent the different General User Interface (GUI) canvases or layers that are present in the currently active mode. Each mode can have multiple SignViews, which again consist of an arbitrary number of SignViewComponents representing the GUI parts of the currently displayed SignView. All SignViews are loaded when a mode becomes active, to reduce loading times between SignView changes. The SignViewComponents are responsible to update GUI information and to react to or propagate user interaction behavior.

Interrupt Handler & Interrupt View

The two remaining types of plugins are the so-called *Interrupt Handlers* and *Interrupt Views*. These plugins differ from the previously introduced ones since they are independent of the currently active mode. Moreover, their main purpose is to actively interrupt the currently running mode and interfering with the currently displayed information. This is done by changing the active view to a so-called Interrupt View, which is displayed for a limited amount of time before it changes back to the previously displayed view. Therefore, Interrupt Handlers are continuously listening for specific events that trigger their execution.

Update Program

The *Update Program* is the second part of MODS and represents the administrative user side of the framework. It runs on devices that are used for maintaining and updating the DS content, like PCs or notebooks and provides functionality for visualizing DS content and modifying it. For this, the Update Program provides a general overview of all the signs that are currently registered. Selected signs can then be displayed in the Update Program as they would be running on the DS device. This can be done since many of the components that are already present in the Sign Program are re-used and extended with so-called *UpdateViews* and *UpdateViewComponents*. This re-use of components enables the Update Program to show and interact with the current content, as it is displayed by the real DS, and, additionally, the extensions provide modification and update capabilities.

For being able to provide both, authentic usability as provided by the DS device and the modification functionalities of the Update Program, it is necessary to distinguish interaction events w.r.t. two different aspects. First, the *context* in which a mode is executed, is used to define where the DS software is currently running. The two possible context options are either the genuine DS device, or the Update Program. The second aspect is the *action* that is conducted to interact with a component. For instance, for the normal interaction with a DS, touch inputs are interpreted as left clicks on specific displayed SignViewComponents. However, in the context of the UpdateProgram additional actions can be defined, which can be leveraged to trigger corresponding update functionalities. Therefore, if a left click is used for normal navigation purposes, a right click could be configured to trigger the implemented update functionality of the SignViewComponent (e.g., opening a picture upload dialogue). This does not have any effect if the software runs in the context of the DS but leads to executing update functionalities in the context of the UpdateProgram.

Office Mode

The *Office Mode* plugin is tailored to represent the use-case of a general office door sign. An exemplary application

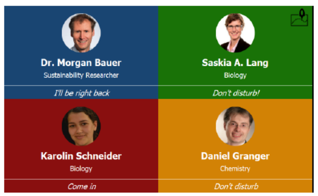


Figure 4: SignView representation

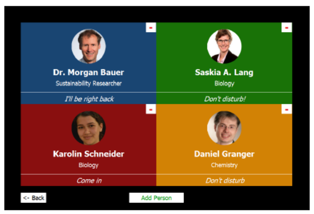


Figure 5: UpdateView representation

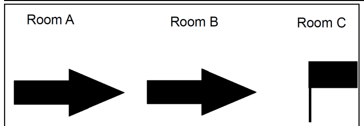


Figure 6: Exemplary symbols displayed after triggering the WayFinder Interrupt Handler. Arrows indicate the suggested walking direction for the users. The flag symbol indicates the goal destination.

for this mode is shown in Fig. 4. This mode provides an overview of the people present in the office, corresponding pictures and naming information. Additionally, a status is provided, which can be modified according to the preferences of the individual people. More detailed information about individual people can be shown by interacting with their profiles. This triggers the navigation towards a more detailed view, which may look as depicted in Fig. 2

Conference Mode
For simple room-booking and calendar functionalities tailored towards seminar and lecture rooms, the *Conference Mode* plugin can be used. Fig. 3 shows the current implementation that allows booking specific time-slots for various dates.

Message Interrupt Handler
The *Message Interrupt Handler* plugin is a prototypical way of communicating between a normal and an administrative user of a sign. In general, this interrupt handler plugin allows an administrative user to send messages to specific signs, which are then displayed for a configurable amount of time. Potential application cases include support-modes that allows administrators to react and directly inform users that retrieve information about problems using DSs.

WayFinder Interrupt Handler
A common problem in foreign environments like new cities or buildings is related to way-finding. The *WayFinder Interrupt Handler* is a plugin that allows a user to query the DS for directions towards, for instance, a specific room in a building, which is also equipped with a DS and selectable from a set of registered devices. The DS then displays one of the symbols shown in Fig. 6, indicating either a movement direction for the user using an arrow symbol or the destination location using the flag symbol. Additionally, the DS returns to its original view after a predefined amount of time and the next DS on the way towards the destined location is invoking the WayFinder Interruption handler to indicate the next steps for the user.

Pre-Study
For assessing the usability of both, the Sign Program and the Update Program a pre-study was conducted. The study encompassed eleven participants (i.e., scientific employees and students) from the domain of computer science. Therefore, a basic understanding of touch-controlled devices could be assumed for all participants. The participants had to fulfill tasks from three main categories: First, the participants had to conduct general navigation and interaction tasks using the DS device. Second, tasks that involved updating displayed information via the Update Program needed to be handled. Finally, the participants had to solve extended update tasks.

Setup
The setup of the evaluation environment consisted of three Raspberry Pi devices acting as DSs and three attached 7" touch screens for user interaction. Additionally, one notebook was set up, which was used to run the database and the Update Program part of the framework. The three Raspberry Pis were configured to represent three door signs in an office environment (i.e., the mode was set to Office Mode). The participants had to conduct 16 different tasks. Those encompassed simple tasks like navigating and displaying information on the DS devices and more sophisticated ones, using the Update Program for updating content, changing modes and update plugins. Afterwards, the participants had to indicate the perceived simplicity for completing the tasks, using a 5-point Likert-scale ranging from $0 - 4$, indicating a task complexity from easy to very difficult.

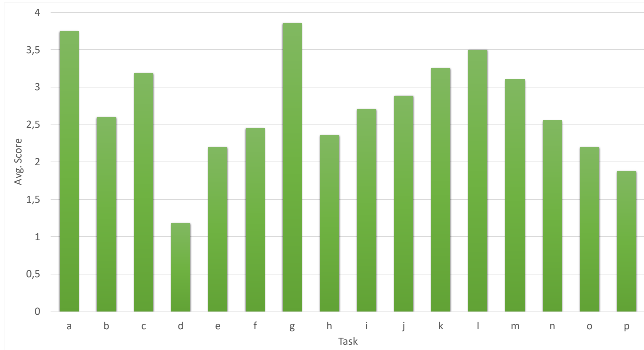


Figure 7: Average ratings of the participants for the 16 conducted tasks labeled as (a-p)

Results

As many people from the domain of computer science are comfortable with using touch-controlled devices, most of the tasks related to general navigation were perceived as easy to perform. The most problems were registered during the interaction with the DS running in the context of the Update Program. Figure 7 shows the average results for the conducted tasks labeled as (a) to (p). The currently implemented *action* for triggering update functionalities (i.e., right-click in contrast to left-click) was perceived as not intuitive and the absence of a functionality outline complicated matters. This is also reflected by the results of the tasks (d,e and f), which all required interacting with plugins using the Update Program. General remarks for improving the work were related to adding visual clues, tool-tips and feedback functionality that informs a user about successfully executing certain functionalities (e.g., successful update).

Conclusion and Future Work

Signage is transitioning from static analogue signs towards Digital Signage (DS) providing a variety of benefits in terms of maintenance, updateability and the support of dynamic content. In this context we presented the Modularly Operated Digital Signage (MODS) framework, facilitating the simple adaption of DS solutions towards diverse use-cases. This is done by organizing use-case specific solutions using plugins, which can then be loaded and orchestrated by the MODS framework. Finally, a pre-study was conducted to receive the first insights into the usability of the MODS. Overall, the usability of the framework was perceived as good, but additional visual clues and information about control functionalities are required to further improve the usability.

As an open and extensible framework, MODS provides a valuable basis for various future developments. Example directions that would especially benefit from the current design encompass context-aware functionalities. One can think of automatic updates regarding presence information, e.g., the digital sign can update itself when a person enters an office (marking the person as “available”) or if the person joins a telephone call (marking the person as “DND”). The way finding module becomes particularly useful if combined with a “meeting organizer” when it enables the routing of external guests. Combined with emerging sensors in the IoT, the usefulness of DS will further increase.

Acknowledgements

This work is based on the Software Campus project “ADRAS“, which is funded by the German Federal Ministry of Education and Research (BMBF) under grant no. “01IS17050‘ and also supported by the German Federal Ministry of Education and Research (BMBF) as well as by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within CRISP.

REFERENCES

1. Colloqi Consulting. 2019. piSignage -Digital Signage for all. (2019). https://www.pisignage.com/, retrieved on 2019-05-17.
2. Goldmedia. 2019. Digital Signage becomes ubiquitous. (2019). https://www.goldmedia.com/en/news/info/article/digital-signage-becomes-ubiquitous/, retrieved on 2019-05-17.
3. John V Harrison and Anna Andrusiewicz. 2004. A virtual marketplace for advertising narrowcast over digital signage networks. *Electronic Commerce Research and Applications* 3, 2 (2004), 163–175.
4. Visix Inc. 2019a. Axis TV Signage Suite. (2019). https://www.visix.com/, retrieved on 2019-05-17.
5. XOGO Inc. 2019b. XOGO Decision Signage - Digital signage made easy. (2019). https://www.xogo.io/, retrieved on 2019-05-17.
6. Jon Knight and Jason Cooper. 2018. Raspberry Pi driven digital signage. (2018).