We should start thinking about Privacy Implications of Sonic Input in Everyday Augmented Reality!

Katrin Wolf[1], Karola Marky[2], Markus Funk[2]

Faculty of Design, Media & Information, HAW Hamburg[1]
Telecooperation Lab, Technische Universität Darmstadt[2]

[1]katrin.wolf@haw-hamburg.de, [2]{surname}@tk.tu-darmstadt.de

Abstract

Evolution in technology causes privacy issues, which are currently under intense discussion. Here, much attention is given to smart cameras, the Internet of Things and the Internet in general, while sonic AR systems are overlooked. Many users, for example, blindfold their laptop cameras with physical layers, but it seems as if no attention is drawn to the sonic hardware that can be hacked just like cameras. In this position paper, we highlight everyday situations that are prone to cause privacy problems through Sonic AR. We then look at current proposals to protect users from camera-caused privacy violations as examples and discuss how they could be adopted to prevent sonic information misuse. We conclude by stating that the current privacy discussion overlooks Sonic AR, although this is a channel across which even more detailed and hence, more sensitive, information can be communicated and misused.

1 Introduction & Background

Augmented Reality (AR) devices are becoming more and more integrated into people's daily lives. While AR solutions using a head-mounted display (HMD) (e.g., the Microsoft HoloLens) are getting ready for the usage outside the lab, many other (also non-visual) AR solutions have hit the market and are used in many everyday situations which leads to many privacy discussions. However, when thinking about AR, many people think about having their picture taken without giving consent, or their picture being streamed to the Internet without their knowledge (Denning et al., 2014). But there is another capability of AR devices, which is

Veröffentlicht durch die Gesellschaft für Informatik e. V. 2018 in
R. Dachselt, G. Weber (Hrsg.):
Mensch und Computer 2018 – Workshopband, 02.–05. September 2018, Dresden.
 https://doi.org/10.18420/muc2018-ws07-0466

becoming more prominent, that has not led to a very prominent privacy debate yet: Voice Input.

When we look at the history of everyday AR devices, the proliferation of Google Glass first seemed to be a technological breakthrough, but then became a concerning topic in many privacy discussions. The public opinion was that persons wearing the head-mounted display were recording passersby without their consent. This led to a shift in the public opinion about HMDs and persons wearing the Google Glass in public were soon called "Glassholes"[1] and eventually led to the Google Glass being discontinued.

Nowadays the debate of wearing HMDs in public has calmed down. Although newer devices e.g. the Microsoft HoloLens also are equipped with cameras, the privacy discussion about these newer devices never started, or not started yet. Have people become sensitized towards head-mounted cameras or did people stop caring about their picture being taken?

The privacy implications of various technologies have been investigated by a plethora of works in different domains. Mobile users consider pictures, videos, their location as well as voice recordings as sensitive data that should be protected (Muslukhov et al., 2012). To address privacy concerns towards devices developers use LED indicator lights. In the case of webcams, indicator lights suffer from a poor effectiveness, because not all users recognize them (Portnoff et al., 2015). Furthermore, the indicator's behavior can be changed by a firmware manipulation (Brocker & Checkoway, 2014) making it untrustworthy. To combat this, Koelle et al. (2018b) provide design requirements that support users in noticing the status of a body-worn camera. Privacy issues do not only concern the primary users, i.e., the users of the AR devices, but also the privacy of secondary users - bystanders - can be compromised (Denning et al., 2014). Many works mention privacy issues based on "recording". This recording does not explicitly exclude voice recording, the main privacy discussion, however, is directed towards video recording.

In this paper, we aim to raise awareness about the voice input and output capabilities of state-of-the-art Augmented Reality devices (Sonic AR) and the non-obvious privacy implications that using these devices comes with.

2 Privacy-Critical Sonic Input Scenarios in AR

Since AR devices have made their way into more and more parts of our daily lives, we provide three example scenarios when voice input can be a privacy issue: *at home*, *at work*, and *in public*.

2.1 At Home

Guests are visiting a user of smart speakers at home. Most hosts fail to tell their guests that a smart speaker (e.g., Google Home or Amazon Alexa) is always listening for voice commands.

[1] https://www.theguardian.com/technology/2014/feb/19/google-glass-advice-smartglasses-glasshole

Although manufacturing companies guarantee that none of the recorded conversations are transferred and stored without a user's explicit input, e.g., a trigger keyword, smart assistants could misinterpret parts of a regular conversation as a voice command and therefore start streaming private conversations to acquaintances[2]. In the future, will we have to ask for consent first when a guest enters our home that is equipped with smart speakers?

2.2 At Work

In most work environments audio privacy is not considered at all. In meetings, smartphones and laptop computers with active microphones are placed directly onto the meeting table. But also, more restricted environments, e.g., production sites of big automotive producers, require visitors to tape their phone cameras before bringing them onto the site. However, visitors are not required to tape their microphones or asked if they activated "always-on" features, i.e., activating the smart phone by a voice command.

2.3 In Public

The third scenario is using voice activated technology in public spaces. Some countries, e.g., Germany, have laws and regulations for filming in public space. When a surveillance camera is active, owners of that camera need to put up a sign warning passerby that their picture might be taken in this area (e.g., BDSG 2018). These laws and regulations apply to voice recordings too but are not that ubiquitously perceivable the way camera warnings are. Therefore, a conversation in a train or on the street could easily be recorded by any of the bystanders' smart devices without the recorded persons' or even without the device owner's knowledge.

3 Proposed Solutions

The permission to take a photograph of a person in the public varies between countries[3]. Photographs potentially enable the identification of depicted persons and are therefore considered as personal data under the EU Data Protection Directive 1995, under its revision from 2018 and under other derived national regulations. Voices also potentially allow identifying depicted persons. Hence, laws for protecting personal rights in this scope, already exist[4], even though nowadays the primary focus is on camera recordings.

Assuming always-on Sonic AR is violating privacy, there are several possibilities for privacy protection. First, always-on Sonic AR could be forbidden, which is not a realistic solution. Second, the usage of Sonic AR could be generally allowed, which seems not appropriate as personal rights, business security, and social concerns would be violated, and it is of personal,

[2] http://www.dw.com/en/amazons-alexa-records-and-shares-private-conversation/a-43924258

[3] Legal regulations defining whether a photograph of a person requires his/her consent vary between locales. An overview is provided at https://commons.wikimedia.org/wiki/Commons:Country_specific_consent_requirements, accessed 14/06/2018

[4] EU GDPD, http://ec.europa.eu/justice/data-protection/reform/index_en.htm, accessed 14/06/2018

industrial, social, and governmental interest to avoid such a right violence. The third possibility is, of cause, to conditionally allow Sonic AR. Similar strategies have been proposed by related research for AR cameras (Koelle et al., 2018b). Hence, we will, inspired by conditional rules allowing for always-on AR cameras, propose ways to transfer promising solutions towards Sonic AR.

3.1 Permission for certain user groups

Applications for special-needs groups, e.g., blind people, are most likely socially accepted, especially if the audio is used for providing in-situ information, such as departure times at train stations, but not for recording (Koelle et al., 2015). Hence, we could imagine that future assistant devices could integrate Sonic AR, which might be indicated as such and follow other laws than the EU Data Protection Directive.

3.2 Embedded signal-based switches or noise

In some secure environments, where industrial innovations are under development, Sonic AR could be forbidden in general. Two possibilities of turning off the Sonic AR devices are possible.

First, the user has the responsibility to switch off the device if required. But users might forget this, because in ubiquitous computing devices become invisible through adaption. Furthermore, users might simply not cooperate. Therefore, the second possibility is controlling the Sonic AR devices by signals. Those signals are given by the ubiquitous environment and can stop the audio processing and/or the audio recording. A malicious Sonic AR device, however, might be capable to continue recording by simply ignoring the signal. Therefore, a third possibility is sending noise in secure environments that project the audio signals, such that they can no longer be processed and interpreted by a computer.

3.3 Mode transparency

The context of the situation, e.g., who we talk to and the conversation's topic, might create the desire that audio/speech is neither captured nor analyzed. An obvious and intuitively User Interface of Sonic AR devices could inform bystanders about the Sonic AR device's status. This would empower the bystanders to refrain from sharing audio/speech information in that situation.

User Interface suggestions from Koelle et al. (2018b) aim to inform bystanders whether a smart camera is processing and/or capturing and/or recording images as well as the application purpose. Similar to that, we suggest to design Sonic AR devices with a User Interface that intuitively and immediately provides information about the device's actions to everybody who might create an audio signal. In doing so, bystanders can choose to either avoid sharing auditory information or to move to a location where their audio/speech is not perceivable by any Sonic AR system.

3.4 Explicit commands to activate

The most limiting Sonic AR User Interface would be a not always-on system, which only switches on if the user takes action. Similarly to the activation mechanism of Google Glass, pushing a button to activate voice input could ensure that the activation is only temporarily and clearly visible for any bystander. Alternatively to gestural input, speech input could be used, such as known from to most state-of-the-art smart assistant devices (“Ok Google”, “Hey Siri”, “Alexa”), although this requires an always active microphone. Both input techniques, gestures and speech commands, may have advantages in being clearly understood and appropriate. While speech may be very clear for indicating in calm environments, such as in a library, speech commands would most probably also disturb bystanders. Pushing a button to activate a voice command might be subtle in quiet environments, but very visible in the user and the bystander are in a face-to-face situation. However, hands-demanding input might be inappropriate in many everyday situations.

4 Conclusion and Outlook

In this position paper, we want to highlight the fact that voice input in everyday Sonic AR environments is treated more and more carelessly with regard to privacy. While video recordings underlie strict laws, regulations, and have social acceptability implications, the society seems to neglect always-on microphones that are required for voice input.

This is why we argue that paying close attention to privacy implications is not only necessary for cameras, but also for microphones. We argue that this is a problem in many everyday situations and provide three example scenarios.

Furthermore, we suggest four different ways of designing voice interaction for Sonic AR more privacy friendly. We hope that designers, stakeholders, and especially users of Sonic AR devices will become more aware of the privacy implications that using such an “always-on” technology comes with.

References

BDSG. (2018). § 4 Bundesdatenschutzgesetz Deutschland.

Brocker, M., & Checkoway, S. (2014). iSeeYou: Disabling the MacBook Webcam Indicator LED. In *USENIX Security Symposium*. USENIX Association. 337-352.

Koelle, M., Kranz, M., & Möller, A. (2015). Don’t look at me that way!: Understanding User Attitudes Towards Data Glasses Usage. In *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI)*. ACM, 362–372. DOI: http://dx.doi.org/10.1145/2785830.2785842

Koelle, M., Boll, S., Olsson, T., Williamson, J., Profita, H., Kane, S., & Mitchell, R. (2018a). (Un) Acceptable!?!: Re-thinking the Social Acceptability of Emerging Technologies. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. W03

Koelle, M., Wolf, K., & Boll, S. (2018b). Beyond LED Status Lights-Design Requirements of Privacy Notices for Body-worn Cameras. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM. 177-187

Muslukhov, I., Boshmaf, Y., Kuo, C., Lester, J., & Beznosov, K. (2012). Understanding users' requirements for data protection in smartphones. In *Proceedings of the International Conference on Data Engineering Workshops (ICDEW).* IEEE. 228-235

Portnoff, R. S., Lee, L. N., Egelman, S., Mishra, P., Leung, D., & Wagner, D. (2015). Somebody's Watching Me?: Assessing the Effectiveness of Webcam Indicator Lights. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM. 1649-1658

Authors

Wolf, Katrin

Katrin Wolf is a professor for Media Informatics at Hamburg University of Applied Sciences. Her research interests lie at the intersection of human–computer interaction and interaction design, focusing on how to make novel technologies more usable and useful. To date, Katrin's research has focused on technologies and domains including mobile and wearable systems; virtual, augmented and mixed reality, as well as interactive exhibitions. She actively volunteers in the HCI research community and, for example, is general chair for MUC 2019 and Student Research Competition chair at CHI 2019.

Marky, Karola

Karola Marky is a doctoral Human-Computer Interaction researcher at Technische Universität Darmstadt. Her fields of expertise are Usable Security and Usable Privacy. So far, Karola's research has focused on the usability of end-to-end verifiable Internet voting schemes and the privacy of mobile device users. Before her master, Karola spent six months as visiting researcher in the "Information and Society" research division at the National Institute of Informatics (NII) in Japan. Karola volunteers in the research community by reviewing papers for conferences, journals and workshops.

Funk, Markus

Markus Funk is a post-doctoral Human-Computer Interaction researcher and area head at the Technical University of Darmstadt, who is an expert in Augmented Reality and Virtual Reality. He holds a PhD in Human-Computer Interaction from the University of Stuttgart. During his PhD, Markus spent a research semester at the Fluid Interfaces group at the MIT Media Lab. Further, Markus is a regular reviewer for peer-reviewed conferences and journals and is a PC member of different academic conferences.