2021 Joint 10th International Conference on Informatics, Electronics & Vision (ICIEV) and 2021 5th International Conference on Imaging, Vision & Pattern Recognition (icIVPR)

2021 Joint 10th International Conference on Informatics, Electronics & Vision (ICIEV) and 2021 5th International Conference on Imaging, Vision & Pattern Recognition (icIVPR) | 978-1-6654-4923-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICIEVICIVPR52578.2021.9564182

Social Augmented Reality: A Multiperspective Survey

Anton Nijholt
University of Twente
Enschede, the Netherlands
a.nijholt@utwente.nl

***Abstract*—We introduce research on social augmented reality with an emphasis on social face-to-face interaction. In social interaction interactants employ knowledge about their conversational partner and have their verbal interaction supported by nonverbal interaction cues. We survey the issues that arise when we pursue social interaction in augmented reality. Since handhelds and bulky head-mounted devices hardly allow unobtrusive interaction we pay extensive attention to developments in the field of smart glasses and smart contact lenses. The focus is on the use of smart augmented reality glasses during social interactions. Acceptance issues and disruption of social interaction due to the use of these devices are also touched upon.**

***Contribution*—An overview of social aspects of interpersonal communication in augmented reality applications.**

Keywords—augmented reality, social interaction, smart glasses, smart contact lenses, social signals, human-computer interaction

I. INTRODUCTION

The usual and often cited definition of augmented reality is that of Ronald T. Azuma [1]. Slightly adapted, it says: Augmented Reality (AR) (1) combines real and virtual objects in a real environment, (2) registers (aligns) real and virtual objects, (3) and runs interactively, in three dimensions, and in real-time. Usually, it is also mentioned that AR should not be limited to specific technologies and that it does not address sight only. AR applies to all senses. We have AR systems that provide visual, auditory, tactile, olfactory, and gustatory experiences or combinations of these experiences. In vision-oriented AR we can distinguish between video see-through, optical see-through (OST), mirror, and projection-based AR. Each of these systems can have stationary and mobile variants. We can also distinguish between head-attached, handheld, monitor-based, and spatial AR.

Some new developments include the embedding of AR, whether it is sight-, sound-, touch-, smell-, or taste-oriented, in the Internet of Things (IoT) (pervasive computing, ambient intelligence), that is, the use of sensors and actuators that enhance the AR experience and that are not intrinsic to the AR technology that is used. Related to this is the emerge of social augmented reality. In augmented reality, we interactively experience a real-world environment that has been augmented with computer-generated perceptual information. This information can range from virtual messages to virtual humans or animals. Moreover, the augmented worlds can be co-inhabited and co-experienced with other humans and by encountering virtual or real humans social interaction should come about naturally. Depending on the kind of AR (optical see-through, video see-through, spatial) and the AR device that is used by the user(s) (handheld, head-attached, device-less), and the computer-generated content (visual, sound, taste, scent, touch, multisensorial), various ways of social interaction and enhancing social interaction can be observed.

In this paper, we have observations on the present AR research and literature that deals with the social interaction aspects of AR applications. From these observations, characteristics of social interaction in AR, augmented social interaction, and social interaction that emerges from and in the context of AR applications are obtained. The focus of this paper is on the use of (optical see-through) smart glasses. We will nevertheless sometimes have observations on Head-Mounted Devices (HMDs) such as Microsoft's HoloLens and others. One cannot expect that these devices will be used in everyday life situations. However, due to the miniaturization of sensors and processing devices, we can expect that many of the capabilities of these bulky and heavy devices will also become available for future versions of smart AR glasses.

Nowadays, AR is called the future of social media. AR pioneer Ronald T. Azuma predicted that AR, in particular optical see-through glasses, will be the dominant platform and interface, supplanting the smartphone, for accessing digital information. Light-weight, non-obtrusive, and optical see-through glasses are seen as "... the best chance of achieving the long-term vision of ubiquitous consumer AR displays." [2].

In the next section, we have some background on AR in the context of social augmented reality. Section III is on social signals in human-human interaction that we would like to preserve for users of augmented reality devices and their social interaction partners. Section IV provides a view on current research on "Digital Selfs", or more general, user profiles that become available to partners in an AR-supported interaction. Section V focuses on AR glasses, their social acceptance issues, and their use in face-to-face conversations. How do they affect the interaction, how do they disrupt natural interaction, how can they be used to enhance social interaction? We will look at future commercial technological developments, in particular as they are planned by Facebook's research. We also take a look at the development of smart contact lenses. Finally, a short section VI contains some conclusions.

II. Background: Towards Social AR

Augmented reality that is shown on smartphones has made the concept popular for a broad audience. Although hard-core AR researchers refuse to call it augmented reality, the Pokémon Go game made consumers aware of the possibilities of integrating virtual and real-world scenes. In the same way, but even simpler, this awareness has been fed with smartphone applications where selfies can be decorated with virtual ears, mustaches, et cetera, the more serious virtual mirror and retail and "magic mirror" applications where users can virtually try on products (clothes, make-up, shoes, glasses) in which they are interested, or the text or audio messages that are added to navigation maps. The general audience has also met AR in outdoor events, for example, in projected AR entertainment applications with interactive animations displayed on walls and buildings where real and virtual objects are integrated into the projection. For these applications no AR HMD or eyewear is necessary.

In contrast, in professional environments AR devices are used for remote collaboration, sometimes requiring bulky and uncomfortable AR HMDS that may weigh more than half a kilo. Less advanced AR is possible with smart AR glasses, also head-up displays, sometimes hooked-up to a smartphone or other handheld that is used as an additional head-down display and control. There is a continuous development of new and smart sensor technology that helps in the further downsizing of equipment while increasing its capabilities. Applications for smartphones will be transferred to smart glasses, applications that are now only possible with advanced (see-through) HMDs will become possible with smart glasses in the future. They will make it possible to have AR applications for daily social life activities.

A rather broad definition of social augmented reality can be found in [3]. Here Social AR is defined as the interactive and social experiences of two or more collocated people enhanced by digital information. Maybe not intended this way, but this definition does not restrict itself to AR enhancements, rather it allows enhancements from whatever sensors and actuators in wearables and in the environments that are available. If we restrict ourselves to interactive and social experiences in physically collocated situations that can be enhanced by AR then we get closer to the social AR view we will explore here.

Who can be our social partners in augmented reality? When we talk about social interaction that involves an AR user, we can distinguish between the interaction that takes place with a non-user, another AR user, an avatar representing another person, a virtual reality (VR) user, or a virtual human. In this paper, we confine ourselves to one or more partners that wear smart glasses. Just a few words about two other important types of interaction that due to space limitations will not be considered here in detail.

The first one is social interaction with remotely located partners. Especially in task-oriented AR, collaboration with remote partners is a well-established application area. Remote partners can be present as avatars on the HMD and the user can see the behavior of the remote partner displayed on the avatar. Social interaction environments are more informal and less task-oriented, but also there we can have AR supported and shared social activity with remote partners, such as enjoying and commenting on a 'joint' meal, discussing a present for a grandchild, commenting on a kid's drawing, playing a game, get help with hanging a painting, get help teaching your child, get help with cooking, deciding about what to buy in a supermarket, have a conversation on what you see during a walk, or sharing verbal and non-verbal enthusiasm when watching, 'shoulder to shoulder', a football match on television.

The second kind of interaction that we will not discuss here is the interaction with virtual embodied characters such as virtual humans, virtual pets, or cartoon characters. They are autonomous virtual anthropomorphic characters that can play the role of assistant or companion in social activities. In [4] we argued that with the disappearing computer [5] virtual humans will be the face of intelligence in our smart environments. In (OST) AR these virtual humans will accompany us or will pop up with task- and context-related knowledge and we can have or develop different types of relationships with them, for example, a domestic servant/employer (butler), buyer/seller, friendship, or love relationship. Presently we see these virtual humans also appear in the AR context [6]. Unlike avatars that are intended to directly depict the behavior of its human owner (obtained from sensors such as cameras, motion captures systems, microphones, et cetera), an 'autonomous' embodied agent such as a virtual human has its perceptive capabilities, behavior, and intelligence arise from algorithms.

The contemporary view of AR is oriented towards the visual perception of the world. In real life, we use all our senses to perceive the world, and usually do this in a multisensorial way, integrating senses that perceive sight, sound, touch, smell, and taste, and our proprioceptive sense that provides information about our body. These other sensory experiences can be augmented as well, but have received limited attention. But there are examples of augmented audio, smell, taste, and touch (haptic) reality as well, often included in a cross-sensory modal point of view, for example, in multisensorial augmented human-food interaction research. Audio in augmented reality addresses the inclusion of audio content in the virtual layer that overlays reality. Just as in "augmented seeing", where we can address incomplete and imperfect visual information, in "augmented hearing" we can consider how to make incomplete or imperfect audio information more audible to the partners in social interaction.

III. Social Signals and AR Interactions

Social signals play an important role in social and other human-human interactions. For example, there is unaware but naturally occurring mimicry when persons interact, listening has many active components, and in (conversational) interactions we become aware of the changing affective and cognitive states of our interaction partner. Is he or she still interested, does he or she feel ill at ease, does our partner look worried or nervous, does he or she speak the truth, how to interpret his smile, etc.? This leads to the question: How does AR affect people's co-located social interactions? And, how can AR enhance people's co-located social interactions?

Natural human-human interaction takes place in the presence of social signals. Technology can disrupt these signals, but technology can also help to enhance and interpret these signals. Which signals are disrupted by which characteristics of AR technology in face-to-face interactions, which signals need to be repaired, and can other capabilities of AR technology help with this? This requires detailed knowledge of the analysis and the synthesis of social signals and their translation to AR (or other) technological means. As will become clear from the next sections, AR technology can not only obstruct natural interaction, it can also enhance co-located social interaction and stimulate interaction among city dwellers during public AR events.

Face-to-face interaction research is usually based on Sachs' turn-taking model of conversations [7]. This concept of prescriptive turn-taking is a kind of pipeline model with a speaker speaking, a listener providing nonverbal feedback, and then speaker and listener switch roles. But within turns, there is also simultaneous expressive behavior. A nice example is mimicking [8] and more generally, expressive feedback that requires synchronization between two or more people interacting with each other. Perception and performance of conversational behavior occur simultaneously and need to be aligned. Many more aspects of human-human, human-animal, or human-agent interaction and activity are based on simultaneity rather than on a sequential model. For example, walking together, shaking hands, performing joint tasks, and cooperative and competitive games and sports activities. In [9] more observations and examples of what we called 'anticipatory synchronization' can be found. In an AR situation, optical see-through devices allow us to walk with a virtual partner, to walk a virtual dog [10], to have musical interaction with real and virtual partners, attend a cocktail party with real and virtual partners [11], to move furniture together and perform similar tasks, that is, social interactions and daily activities where coordination by mutual adjustments is needed.

How can AR support conversational interactions in general and support interaction activities that require a mechanism of mutual adjustments in particular? We can ask AR researchers to take up the challenge of using AR to support rather than disrupt affective and social human-human interaction behavior and other synchronized joint human activity, including human-agent interaction behavior and activity.

What do we know about our conversational partners? What information is digitally available? As an example, Facebook knows about us and our Facebook friends. AR interaction with these friends, whether this knowledge is obtained from Facebook, the many other social media, or otherwise collected from websites or publicly available or hacked databases, can be supported by their profile information. It can be displayed in the AR-created virtual layer that is superimposed on reality. Some modest attempts can be found in the next section. There we will also address how to superimpose and use profile information in a visual layer associated with our conversational partner. AR devices can have sensors that are meant to capture affective information from conversational partners. When wearing (optical see-through) glasses in face-to-face conversations the lower face region remains visible. However, eye and eyebrow movements will be difficult to detect. Cameras or other sensors integrated with the AR device make it nevertheless possible to capture this information and make it available to one or both conversational partners [12]. Rather than have such information captured from your partner's head-mounted capturing device and have it sent to your device for interpretation, facial expression understanding can be included in your own device. Sometimes it can be useful to know what affective information is given away by your facial expressions. Sometimes it is useful to obtain support in understanding someone's facial expression behavior. Did the system recognize your partner's smile as a genuine (Duchenne) smile or just a polite, fake smile?

There are other ways to measure stress, cognitive load, or emotions with body-worn sensors and to factor those results into the use of an AR device. Brain activity can be distinguished with EEG scalp or ear sensors. Glasses have been designed where emotions of the wearer are inferred from camera-captured local facial expressions (around the eye), skin conductance of the nose, and pulse from an earlobe [13]. Emotional and cognitive state monitoring with pressure sensors in a headband has also been explored [14]. In other words, AR eyewear can use techniques other than computer vision to monitor the cognitive or emotional state of the user and make it available to the user and its interaction partners. Real-time obtained information can also be integrated into profiles of interaction partners and be used in future interactions.

IV. "DIGITAL SELFS" AND "DIGITAL OTHERS"

In the 1901 book "The Master Key: An Electrical Fairy Tale." by L. Frank Baum [15] Robert, the protagonist, is offered a "character marker". What is a Character Marker? As explained in the book, people try to appear good when they are not, they seem to be friendly and kind when in reality they are not. Hypocrisy is a human failing. Hence, the Character Marker. "It consists of this pair of spectacles. While you wear them every one you meet will be marked upon the forehead with a letter indicating his or her character. The good will bear the letter 'G', the evil the letter 'E'. The wise will be marked with a 'W' and the foolish with an 'F'. The kind will show a 'K' upon their foreheads and the cruel a letter 'C'. Thus you may determine by a single look the true natures of all those you encounter." Robert does not want to use the spectacles on his family members: "At one time he thought of putting on the electric spectacles and seeing what the real character of each member of his family might be; but a sudden fear took possession of him that he might regret the act forever afterward. They were his nearest and dearest friends on earth, and in his boyish heart he loved them all and believed in their goodness and sincerity. The possibility of finding a bad character mark on any of their familiar faces made him shudder, and he determined then and there never to use the spectacles to view the face of a friend or relative. Had any one, at that moment, been gazing at Rob through the lenses of the wonderful Character Marker, I am sure a big "W" would have been found upon the boy's forehead."

This 'character marker' can be seen as an example of augmented reality made possible by smart glasses. AR devices offer the possibility to display personality characteristics on or around the head with the person we want to talk to. When our AR device has access to information about our interaction partner, this information can be displayed on our smartphone or our AR glasses and it can be used to have a more satisfying and profitable interaction. Can we really expect to gather the personality characteristics of users of AR devices and make these characteristics available to conversational partners?

A. Profile and Real-time Behavior Information for AR

We should distinguish between information that has been collected in the past and that has been included in an interaction partner's profile and information that becomes available during the real-time interaction. The former can, in Facebook terms, be made available for some friends, all friends, groups, or for the general public. The latter is about the information that emerges during the interaction and is about an interaction participant's verbal and nonverbal behavior during the interaction.

One way to obtain information about our interaction partner is by having him or her identified by face recognition. Many of our friends' faces are on Facebook or other social media. Companies have facebooks for their employees. Faces and associated information about their owners can be collected from social media. When our AR device notices that we want to start an interaction, it can provide us with useful information about our interaction partner, ranging from personal facts, personality characteristics, preferences, and opinions, to mood and emotion. Real-time info has to be interpreted in the context of the already available information and be integrated with it: who is the user, what is the task, what is the location, preferences that are known, physical characteristics, background, and social media activity.

While interacting we can be informed about matters that are currently at play with our conversational partner. What can we learn about the user when we can observe his or her behavior during some period? What can we learn from behavioral information captured by the sensors in AR devices? In [16] results are reported from short observations of expressive behavior. Personality judgments from "thin slices of behavior" are also discussed in [17] and [18]. These short-time observations provide real-time information about a partner that we have not met before and algorithmic interpretation of that behavior can also be added to the profile of a conversational partner that we can access and maintain by updating. Information about an AR user can also be obtained from messages or objects that the user leaves in his AR environment for his friends and others, that is, in a shared augmented reality.

Rather than having profiles and matches generate from social media or obtained from interactions, in [19], the authors discuss the use of user-generated digital profiles or "Digital Selfs". The content of a "Digital Self" is likely to be different from the content that can be generated from databases maintained by all kinds of commercial and (non-)governmental organizations. Moreover, when looking at social interactions augmented by AR the choice can be made to focus not only on identifying and fostering similarities between interests, useful for small talk and breaking the ice but also on dissimilarities that can support opinion-changing discussions. The approach was evaluated as socially acceptable by the participants of their experiments.

It should be noted that collected information about a user can be useful for the user as well. He may adapt his behavior when confronted with such personal behavior information. How does he want to present himself to his interaction partners? Who owns a user's identity and is allowed to use and communicate it? Contemporary social AR applications are still a long way from the scenarios sketched in this section. Nevertheless, they provide us with insight into the future use of AR devices, including HMDs and smart glasses.

B. Social Networks and Social Apps for AR

For completeness, we mention some rather primitive AR that is used on social apps like Facebook, Instagram, TikTok, Pinterest, and Snapchat. Rather than allowing 'full' AR, these apps do add simple AR effects to pictures that can be shared with friends. An extension that not only augments your social media presence with personal information but also makes it available to friends that you meet in real life is the Octi [20] social network. Octi has the faces of everyone in the network available in its databases. With Octi you see one of your network friends in real life, you point your smartphone's camera at her and after she has been identified with facial recognition software she will be surrounded by a floating belt of personal virtual items such as favorite songs and photos. As mentioned by a social media analyst " ... [it] gives teens a compelling reason to be present and communicate with their phones, while gathered with their closest friends." And, " ... It turns you into a walking social media profile." It is an example of a simple augmented reality application on a smartphone that augments real-life face-to-face interaction between social media friends.

Face, body pose, and emotion recognition are some of the technologies that find their way into these commercial social networks. In [21] photos and other information of a user's Facebook friends have been collected in an offline stage. A mobile phone can be pointed to a person to capture his face. With face recognition, running on a server, the system recognizes the face and a visual tracker on the phone overlays information about the person on the screen.

C. Social Networks and AR HMDs and Glasses

Presenting useful information during face-to-face conversations with an HMD with eye-tracking is presented in [22]. Information boxes appear around a conversational partner's head. The boxes fade in or out depending on the visual attention that is given to them. Their opacity changes. Another example of AR face augmentation is HoloFace [23]. A frontal camera of a HoloLens device is used to detect the face and facial attributes, followed by 3D head pose estimation. Items such as objects or animations can be rendered on, or around the subject's face and head. For example, animations can be triggered when the subject opens his mouth or smiles. A natural extension could consist of facial and thus identity recognition for personalizing virtual objects and animations.

In the following examples, smart glasses are used as AR devices. Talk2Me [24] is a suggestion for an AR social network that aims at facilitating social interaction with physically nearby persons. In this network, a user can broadcast messages with an embedded face-signature to nearby persons. With facial recognition, the person can be recognized and further information can be shared in person. Hence, there is no need that people already know each other. Users have to look around if they wear smart glasses or have to point their wearable camera-enabled device to other people's faces to get a match between signature and face. An AR-based chat messenger for smart glasses is ‘AR Emotional Messenger’ [25]. It allows multiple users represented as virtual characters to communicate in 3D chat rooms. Emojis with effects can be chosen to express emotions. A more elaborate augmentation of face-to-face interaction is presented in [26]. Here, shared interests between individuals were identified by finding matches in their profiles as presented in LinkedIn. Using recommender systems techniques, topics also received a ranking. To support a conversation, personalized topic suggestions were projected onto Google's smart glasses. During conversations there can be transitions from different kinds of topics, for example, starting with a ‘safe’ topic as a conversation-opener and going to topics that are more intimate later. Also, while in a conversation during a multi-user setting your glasses can help you to find someone else to talk with. Therefore, showing the right suggestions at the right moment is an important issue. Moreover, different personalities can show different appreciation of such a system. The authors mention the difference in appreciation between introvert and extrovert personalities.

V. AR Glasses, Smart Lenses and Social Interaction

Technology development since 2012, the boosted interest in augmented reality, the assumption to be able to deal with privacy issues, and the assumed interest of consumers in new gadgets, has led to a renewed interest of companies to develop smart glasses for the consumer market and of start-up companies to develop applications for the consumer market. Rather than having two or more persons interact with each other, talking with each other with their smartphones at hand for use, we can expect to see two or more persons interact with each other wearing smart glasses that provide them with auditory and visual support on an individual or shared basis. For that reason, in this section, we zoom in on the use of smart glasses for augmented reality. This section concludes with a short subsection on the development of smart contact lenses for AR.

A. Google Glass: Acceptance and Interaction Use

Google Glass was introduced in 2013, a wearable heads-up display computer resembling regular eyeglasses with a touchpad, a camera, GPS, microphones, and optical see-through AR glasses through which a user can see virtual content overlaid on the real world. And, of course, a processor, a memory, and an Internet connection. It allowed you to take a photograph by winking or reading your email during breastfeeding [27]. It also generated controversy. Do you want someone sitting across from you to take a picture or do a video recording while you have a conversation? Or have the glasses’ camera identify you and then search the Internet for more information about you on social media? Or your facial expressions to be recorded and interpreted, and have your conversational partner access to this interpretation? As an example, in Fraunhofer's SHORE project [28] it was investigated how Google Glass could be used to read your partner’s emotions (Fig. 1). Facial recognition software was used to recognize in real-time the age and gender of a conversational partner, and affective information such as angriness, happiness, surprise, and sadness. Google Glass was banned from bars, cinemas, and casinos because of the discomfort it caused among visitors and fear of violating privacy laws. The growing resentment against users led to them being called “glassholes” or zombies that were ‘glassed out’ during their interaction with others. Google even had to introduce restrictions to keep porn apps from the Internet-linked glasses. All this usage and potential usage triggered so much anxiety concerning privacy and breach of social norms that it was decided to discontinue the glasses in their then form for the consumer market. Rather interest turned to the use of smart glasses in the workplace (Google’s “Glass at Work program”), with applications where privacy is not a concern.

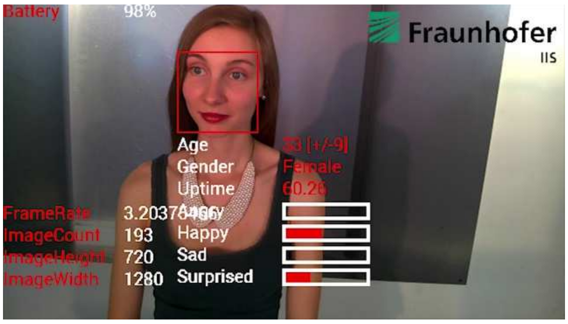

Figure 1. Google Glass with face recognition

As a consequence, there is not much research to report on social interaction with smart glasses. In [29] some observations on face-to-face meetings with one Google Glass user and two or more non-users are presented. How will the use of glasses interfere with the turn-taking process? In a sense, Google Glass can be considered as a non-human participant that takes turns with its user but interrupts the ongoing social interaction between the participants. The non-user co-participants do not necessarily make sense of the actions performed by the glass user and see their interaction interrupted and suspended. However, one can argue that with more advanced AR and processing technology now available, the desired content will be available more quickly and can be displayed more adequately. The need for certain information can be predicted and it can be pushed, rather than searched for during an interaction.

Being simultaneously engaged in another activity during a conversation is often hardly a problem. We can have a conversation while driving, cooking, eating, walking, or playing. Sub-activities of such a multiactivity can progress in parallel or sequentially without disturbing the flow of exchange of social signals. Attention-drawing events that disrupt the flow can occur, unexpected behavior of a car in front, a loud ringtone on a smartphone, or a doorbell ringing. However, paying attention to your smart glasses can be more disruptive than these events and more disruptive than consulting a smartphone or smartwatch because the cause of the disruption in the flow of social signals is not immediately perceptible to the user’s conversational partners. The user takes a turn with a partner that is invisible to his human partners. A comparative study [30] on making profile information available on a smartphone, a smartwatch, or smart glasses concluded that for support of conversations, participants preferred a smartphone or a smartwatch that could be used for screen sharing, while from a privacy point of view the preference was given to the glasses. It should be noted, however, that this study concerned profile information only.

In [31] smart glasses are studied from a technology acceptance point of view. They notice that glasses are for consumers to wear in social environments and that makes that wearables such as glasses are not only perceived as technology but also as a form of fashion. Therefore they argue that theories that explain fashion adoption and use also play a role in explaining smart glasses adoption. Hence, the usual factors in technology acceptance research (perceived ease of use and perceived usefulness) should be extended with factors such as hedonic and social benefits. Their study with questionnaires among a student population has the conclusion that " ... in order to understand smart glasses more fully, scholars and managers need to think of them in terms of fashnology rather than just as a novel technology". Here "fashnology" represents consumer perceptions of wearable technology as a combination of "fashion" and "technology". In support of their conclusion, for the initial Google Glass Diana von Furstenberg-designed versions were available. Now, in 2021, Facebook is teaming up with Ray-Ban to make stylish designs for their future smart glasses available. In a different study on social acceptance [32] aesthetics was not found to be an important factor, though it was mentioned that increased familiarity and usability will probably lead to more impact on acceptance. Social acceptability has also been investigated in [33]. In their research, they look at many different scenarios of smart glasses usage. One of their conclusions is that a negative attitude might diminish over time. Discomfort can disappear through increased familiarity and decreased novelty, the emergence of a new etiquette, and education. We think that is what will happen now and this will certainly be stimulated by having lower prices for consumers. So, it's time again to look at the consequences of using smart glasses during daily activities, at home and in public spaces, and for social interactions.

Presently many other companies than Google have entered the smart glasses competition and they now usually emphasize privacy aspects of their research to see smart glasses enter a mass market of social media usage, rather than having the use of smart glasses restricted to professional communities. We now have Google Glass, Amazon's and Vuzix's glasses, Facebook's ARIA research smart glasses (see the next sub-section), and the announcement of Apple's augmented reality glasses. Besides, we can mention Microsoft's interest in optical see-through AR, for example, its research on HoloLens applications in the Holoportation project [34].

B. *Facebook's Augmented Reality Glasses Research*

In Facebook Reality Labs (FRL) AR glasses have been adopted as a research device. In their ARIA research project, announced in 2020, AR glasses are developed that can be considered as all-day wearable computers with an expected weight of fewer than 70 grams (see Fig. 2). However, they are much more than the usual processors of information. The glasses include cameras to visually observe the world and do head, hand, and object tracking, there are inward-facing cameras to monitor the direction of the wearer's gaze, and seven embedded directional microphones that take care of spatialized audio capturing. The project includes research on interfaces such as EMG wrist sensors and EEG devices that issue neural commands to manipulate objects, beamforming audio to eliminate background noise, lifelike avatars that represent human interaction partners, and the building of LiveMaps, 3D representations of parts of your real-world in which you can share information and virtual objects with your friends and others. Information about the user of the ARIA glasses is captured and can be shared, for example with Facebook friends.

In this ARIA project, a comprehensive view on how augmented reality, including sound, may play a role in future, digitally augmented social interaction. The project includes research on augmented reality (social) interactions, capturing humans and human interaction, modeling and interpreting urban environments from an egocentric point of view, and, in particular, capturing and enhancing speech interaction between users in co-located environments.

All of this is planned to become possible with the future Facebook Glass device that not only provides a view on virtual visual objects added to a virtual layer overlaid on reality but also the introduction of virtual auditive objects or enhancements of existing, naturally occurring audio and speech events in the user's augmented environment. With audio beamforming voices are extracted from a soundscape, in-ear monitors can filter and mute ambient sound, and the microphones can be adapted to the acoustic characteristics of a user's ears. Virtual conversations can be made indistinguishable from real face-to-face conversations, and in a real face-to-face conversation in a noisy environment, your glasses will pick out your conversational partner's voice and enable you to have a relaxed conversation.

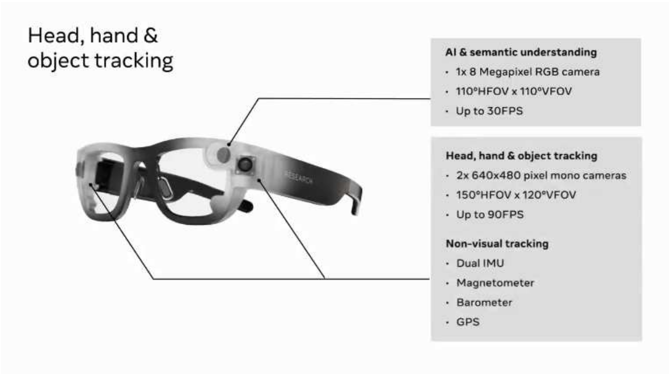

Figure 2. Google's research tool for smart glasses

A user of the ARIA glasses will observe and help to model urban environments, providing an egocentric view while surveying the environment. Capturing humans and their nonverbal behavior and make this information available to others, for example, for driving an avatar in a remote location, is another issue. From what we discussed above, there are many possibilities to use and enhance captured digital social interaction signals, complete or incomplete, verbal or nonverbal. The removal of objects, for example, those that occlude part of the face, and the introduction of other virtual objects in the user's virtual environment, has received attention as well. But having such views integrated with auditory or other (multi) sensorial has not or has hardly received attention.

C. *Eye-mounted Displays and Beyond*

Even less obtrusive and conspicuous during social interaction than glasses are AR contact lenses (ARCLs), sometimes referred to as eye-mounted displays. They can be considered as the ultimate heads-up display. Various companies such as Mojo Vision, Innovega, and InWith, sometimes supported by traditional contact lenses companies, are developing lenses that will contain microLED displays, ultra-thin image and motion sensors, and batteries that draw their power wirelessly from a smartphone's battery or the natural blink process of the eyes. Not only the more traditional contact lens companies but also Google has been researching the inclusion of microcontroller chips and ultra-thin sensors that measure glucose levels in tears to assist people with diabetes. In 2015 Samsung filed a patent for ARCLs (granted in 2019) with a display unit, thin-film camera, and a motion sensor, wirelessly connected to a smartphone. In Samsung's patent [36] eyeball focus and blinks operate the in-built camera. Virtual content floats in front of the user's eyes. It is supposed to consist of text messages or textual information aligned with objects and locations in the real world and in the field of view of the wearer.

Similar views on ARCLs exist among Mojo Vision researchers [37] (see Fig. 3). In their plans, in addition to the lenses, the user has an Internet-connected wearable processor that stores information, wirelessly communicates with the lenses and their sensors, and computes the next frame with virtual content that will displayed to the wearer. For virtual content, one should think of text with weather and traffic information rather than virtual 3D objects. However, their starting point is to provide people with vision impairments with vision enhancement by doing real-time edge detection of objects and augmenting them by highlighting their contours. This makes it more easy to quickly recognize a situation and react to it.

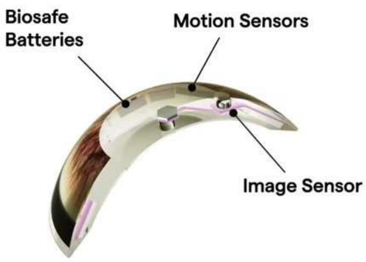

Figure 3. Smart contact lenses (Mojo Vision)

Lenses must be custom-made and require a medical prescription with information about visual acuity, the shape of the cornea, and the size of the iris and pupil. They are classified as wearable medical devices and development, marketing and application must comply with guidelines concerning health and safety. This will be even more the case with intraocular lenses, retinal implants, or visual cortex implants [38]. Apart from appearances in science-fiction films, we cannot expect these devices to enter the consumer market within the next ten years.

VI. CONCLUSIONS

Until now, AR research has focused on AR technology and its possible applications in task-oriented environments. Remote collaboration has been another important issue. Many applications have been investigated in retail and marketing, training, and education. But AR has also become part of entertainment, in theatres, stage performances, location-based games, and public events. Research groups have experimented with entertaining public spatial AR involving computer vision, animations, and interactive play on façades. Handheld AR applications are booming. They include augmented social interaction in which interaction partners can profit from visual and auditory information that otherwise remains hidden. Augmented and virtual events can be shared, manipulated, and discussed.

Traditional AR is aimed at how to register the virtual objects in the world, the interaction between the real and the virtual objects, and the user’s view of the augmented world. Advances in technology coupled with lower prices allow wider use of AR HMDs. The first attempts to introduce AR glasses on the consumer market were not very successful yet. More than before, companies are now aware of the demands placed on privacy by governments and the public. New interest in AR glasses is emerging. Well-known technology companies such as Facebook, Google, Apple, and Microsoft are moving into this market, supported by extensive internal research. As we noted in the Introduction of this paper, AR pioneer Ronald T. Azuma mentioned that light-weight, non-obtrusive, and optical see-through glasses can be seen as the best chance of achieving the long-term vision of ubiquitous consumer AR displays [2], that is, enter a world of ‘ever-present AR’.

In this paper, we surveyed the many aspects that come into play as we move towards large-scale AR use, and in particular, AR use in the social day-to-day interaction with each other. Two comments can be made here. First, there may also be situations in a future social context where we want to use advanced HMDs with additional, advanced capabilities that we will not immediately see realized with lightweight, non-obtrusive eyewear. Think of game situations, amusement parks, and home situations where we accept that we are less mobile and can also offset other disadvantages against the advantages. Second, the core of AR research necessarily focused on the development of AR technology for the individual user and the use of AR technology for remote collaboration between two users. We can now pay more attention to the development of AR and AR applications that are embedded in smart environments (pervasive computing) and the Internet of Things, where the user with all their wearables is also a node in the IoT. The real world that is augmented is already smart. In recent papers, we see this awareness gaining attention, for example in [39], a review on the convergence of AR, intelligent agents, the IoT, and AI, in [40], a survey on the integration of AR and IoT, and a review in [41] on the combination of AR and IoT. It shows that the use of AR will be integrated with our daily activities, including our social activities, in an already smart world. Finally, the environment and the IoT can be given more decision-making powers over the use of AR and the virtual content and decrease therefore the decision-making options of the AR and its user.

REFERENCES

[1] R.T. Azuma, “A Survey of Augmented Reality,” Presence: Teleoperators and Virtual Environments 6:4, pp. 355-385, 1997.

[2] R.T. Azuma, “The road to ubiquitous consumer augmented reality systems,” Hum Behavior & Emerging Technologies 1, pp. 26–32, 2019.

[3] I. Hirskyj-Douglas, A. Kantosalo, A. Monroy-Hernández, J. Zimmermann, M. Nebling, M. Gonzalez-Franco, “Social AR: Reimagining and Interrogating the Role of Augmented Reality in Face to Face Social Interactions,” The Future of Social AR, workshop at 23rd ACM Conf. of Computer-Supported Collaborative Work and Social Computing, 2020.

[4] A. Nijholt, “Where computers disappear, virtual humans appear,” Computers & Graphics 28(4), pp. 467-476, 2004, doi: 10.1016/j.cag.2004.04.002

[5] M. Weiser, “The computer for the 21st century,” Scientific American, 265(3), pp. 94–104, 1991.

[6] A. Nijholt, “Augmented Reality Humans: Towards Multisensorial Awareness,” Proceedings 6th International Conference on Digital Economy (ICDEc 2021), R. Jallouli, M.A.B. Tobji, H. Mcheick, G. Piho, Eds., LNBIP. Cham, Switzerland: Springer, 2021, to appear.

[7] H. Sacks, E.A. Schegloff, and G. Jefferson, “A simplest systematics for the organization of turn-taking for conversation,” Language 50(4), pp. 696-735, 1974.

[8] X. Sun, J. Lichtenauer, M. Valstar, A. Nijholt, and M. Pantic, “A Multimodal Database for Mimicry Analysis,” Affective Computing and Intelligent Interaction (ACII 2011), S. D’Mello et al., Eds., LNCS, vol 6974. Berlin, Germany: Springer, 2011, pp. 367-376.

[9] A. Nijholt, D. Reidsma, H. van Welbergen, R. op den Akker, and Z. Ruttkay, “Mutually Coordinated Anticipatory Multimodal Interaction,” Verbal and Nonverbal Features of Human-Human and Human-Machine Interaction. A. Esposito et al. Eds., LNCS, vol 5042. Berlin, Germany: Springer, 2008, pp. 70-89, doi: 10.1007/978-3-540-70872-8_6

[10] N. Norouzi, K. Kim, M. Lee, R. Schubert, A. Erickson, J.N. Bailenson, G. Bruder, and G. Welch, “Walking Your Virtual Dog: Analysis of Awareness and Proxemics with Simulated Support Animals in Augmented Reality,” Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2019, pp. 253-264.

[11] M.R. Miller, H. Jun, F. Herrera, J.Y. Villa, G. Welch, and J.N. Bailenson, “Social interaction in augmented reality,” PLOS ONE 14(5): e0216290, 2019, doi: 10.1371/journal.pone.0216290

[12] A. Nijholt, “Capturing Obstructed Nonverbal Cues in Augmented Reality Interactions: A Short Survey,” Proceedings of International Conference on Industrial Instrumentation and Control - I2C 2021, S. Bhaumik, S. Chattopadhyay, T. Chattopadhyay, S. Bhattacharya, Eds., Lecture Notes in Electrical Engineering. Cham, Switzerland: Springer, 2021, to appear.

[13] J. Kwon and L. Kim, “Emotion recognition using a glasses-type wearable device via multi-channel facial responses,” ArXiv, abs/1905.05360, Preprint, 2019.

[14] B. Zhou, T. Ghose, and P. Lukowicz, “Expressure: Detect Expressions Related to Emotional and Cognitive Activities Using Forehead Textile Pressure Mechanomyography,” Sensors, 20(3):730, 2020, doi: 10.3390/s20030730

[15] L.F. Baum, “The Master Key: An Electrical Fairy Tale,” Indianapolis, USA: The Bowen-Merrill Company Publishers, 1901.

[16] N. Ambady and R. Rosenthal, “Thin slices of expressive behavior as predictors of interpersonal consequences,” Psychological Bulletin, 111:2, pp. 256-274, 1992.

[17] P. Borkenau, N. Mauer, R. Riemann, F.M. Spinath, and A. Angleitner, “Thin Slices of Behavior as Cues of Personality and Intelligence,” Journal of Personality and Social Psychology. 86(4), pp. 599-614, 2004.

[18] A. Nijholt, “Capturing Immediate Interests in Ambient Intelligence Environments,” Proceedings IADIS International Conference on Intelligent Systems and Agents (ISA 2007), A. Palma dos Reis et al., Eds. Lisbon, Portugal: IADIS Press, 2007, pp. 91-98.

[19] M. Kytö and D. McGookin, “Investigating user generated presentations of self in face-to-face interaction between strangers,” International Journal of Human Computer Studies 104, pp. 1-15, 2017.

[20] J. Fuisz, "Octi: The First People-powered Social AR Platform," https://www.octi.tv/, 2020.

[21] M. Dantone, L. Bossard, T. Quack, and L. van Gool, “Augmented faces,” IEEE Conference on Computer Vision Workshops, 2011, pp. 24–31.

[22] R. Rivu, Y. Abdrabou, K. Pfeuffer, A. Esteves, S. Meitner, and F. Alt, “StARe: Gaze-Assisted Face-to-Face Communication in Augmented Reality,” ACM Symposium on Eye Tracking Research and Applications (ETRA ’20 Adjunct). New York, USA: ACM, Article 14, 2020, pp. 1-5.

[23] M. Kowalski, Z. Nasarzewski, G. Galinski, and P. Garbat, “HoloFace: Augmenting Human-to-Human Interactions on HoloLens,” 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, 2018, pp. 141-149, doi: 10.1109/WACV.2018.00022

[24] J. Shu, S. Kosta, R. Zheng, P. Hui, "Talk2Me: A Framework for Device-to-Device Augmented Reality Social Network," IEEE International Conference on Pervasive Computing and Communications (PerCom), Athens, 2018, pp. 1-10.

[25] J. Choe, T. Lee, S. Seo, "Augmented-Reality-Based 3D Emotional Messenger for Dynamic User Communication with Smart Devices," Electronics 9, 1127, 2020.

[26] T.T. Nguyen, D.T. Nguyen, S.T. Iqbal, and E. Ofek, “The known stranger: supporting conversations between strangers with personalized topic suggestions,” Proceedings 33rd Annual ACM Conf. on Human Factors in Computing Systems. New York, USA: ACM, 2015, pp. 555–564.

[27] D. Schuster, “The revolt against Google ‘Glassholes’,” New York Post, July 14, 2014, https://nypost.com/2014/07/14/is-google-glass-cool-or-just-plain-creepy/

[28] SHORE https://www.youtube.com/watch?v=Suc5B79qjfE, 2013

[29] B. Due, “The social construction of a Glasshole: Google Glass and multiactivity in social interaction,” PsychNology Journal, 13(2-3), pp. 149 – 178, 2015.

[30] I. Hirskyj-Douglas, M. Kytö, and D. McGookin, “Head-mounted Displays, Smartphones, or Smartwatches? -- Augmenting Conversations with Digital Representation of Self,” Proc. ACM Hum.-Comput. Interact. 3, CSCW, Article 179, 2019, pp. 1-32, doi: 10.1145/3359281

[31] P.A. Rauschnabel, D.W.E. Hein, J. He, Y.K. Ro, S. Rawashdeh, and B. Krulikowski, “Fashion or Technology? A Fashnology Perspective on the Perception and Adoption of Augmented Reality Smart Glasses,” i-Com, 15(2), 2016, pp. 179-194, doi: 10.1515/icom-2016-0021

[32] S. Kernaghan, “Google Glass: An Evaluation of Social Acceptance,” Thesis, School of Engineering and Digital Arts, University of Kent, 2016.

[33] M. Koelle, M. Kranz, and A. Möller, “Don’t look at me that way! Understanding User Attitudes Towards Data Glasses Usage,” in Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI ’15). New York, USA: ACM, 2015, pp. 362–372, doi: 10.1145/2785830.2785842

[34] S. Orts-Escolano et al., “Holoportation: Virtual 3D Teleportation in Real-time,” Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST ’16). New York, USA: ACM, 2016, pp. 741–754, doi: 10.1145/2984511.2984517

[35] M. Abrash, "Project ARIA. In: Facebook Connect Keynote: Oculus Quest 2, Project Aria & More! 1.33.20 - 1.47.50, 2020. https://www.youtube.com/watch?v=woXmJMw2lTM

[36] Samsung Electronics Co., Ltd. “Smart contact lenses for augmented reality and methods of manufacturing and operating the same,” US 2016/0091737 A1. 2015 (Granted in 2019).

[37] T.S. Perry, “Augmented Reality in a Contact Lens: It’s the Real Deal,” IEEE Spectrum. [Online, 16 January 2020] https://spectrum.ieee.org/view-from-the-valley/consumer-electronics/portable-devices/ar-in-a-contact-lens-its-the-real-deal

[38] B.C. Kress, “Smart Contact Lenses and Beyond,” Chapter 17 in: Optical Architectures for Augmented-, Virtual-, and Mixed-Reality Headsets, pp. 185-190. SPIE Press, Bellingham, WA, USA, 2020.

[39] N. Norouzi, G. Bruder, B. Belna, S. Mutter, D. Turgut, and G.A. Welch, “A Systematic Review of the Convergence of Augmented Reality, Intelligent Virtual Agents, and the Internet of Things, Artificial Intelligence in IoT,” F. Al-Turjman, Ed. Cham, Switzerland: Springer, pp. 1–24, 2019, doi:10.1007/978-3-030-04110-6_1

[40] D. Jo and G.J. Kim, “AR Enabled IoT for a Smart and Interactive Environment: A Survey and Future Directions,” Sensors, 19(19):4330, 2019, doi: 10.3390/s19194330

[41] J.C. Kim, T.H. Laine, and C. Åhlund, “Multimodal Interaction Systems Based on Internet of Things and Augmented Reality: A Systematic Literature Review,” Appl. Sci., 11, 1738, 2021, pp. 1-33.