Association for Information Systems
AIS Electronic Library (AISeL)

Research Papers ECIS 2019 Proceedings

5-15-2019

HOW TO INTERACT WITH AR HEAD MOUNTED DEVICES IN CARE WORK? A STUDY COMPARING HANDHELD TOUCH (HANDS-ON) AND GESTURE (HANDS-FREE) INTERACTION

Michael Prilla
Clausthal University of Technology, michael.prilla@tu-clausthal.de

Marc Janßen
Clausthal University of Technology, marc.janssen@tu-clausthal.de

Timo Kunzendorff
Clausthal University of Technology, timo.kunzendorff@tu-clausthal.de

Follow this and additional works at: https://aisel.aisnet.org/ecis2019_rp

Recommended Citation

Prilla, Michael; Janßen, Marc; and Kunzendorff, Timo, (2019). "HOW TO INTERACT WITH AR HEAD MOUNTED DEVICES IN CARE WORK? A STUDY COMPARING HANDHELD TOUCH (HANDS-ON) AND GESTURE (HANDS-FREE) INTERACTION". In Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden, June 8-14, 2019. ISBN 978-1-7336325-0-8 Research Papers.
https://aisel.aisnet.org/ecis2019_rp/76

This material is brought to you by the ECIS 2019 Proceedings at AIS Electronic Library (AISeL). It has been accepted for inclusion in Research Papers by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

HOW TO INTERACT WITH AR HEAD MOUNTED DEVICES IN CARE WORK?

A study comparing Handheld Touch (hands-on) and Gesture (hands-free) Interaction

Research paper

Prilla, Michael, Human Centered Information Systems, Clausthal University of Technology, Clausthal-Zellerfeld, Germany, michael.prilla@tu-clausthal.de

Janßen, Marc, Human Centered Information Systems, Clausthal University of Technology, Clausthal-Zellerfeld, Germany, marc.janssen@tu-clausthal.de

Kunzendorff, Timo, Simplifier AG, Würzburg, Germany and Human Centered Information Systems, Clausthal University of Technology, Clausthal-Zellerfeld, Germany, timo.kunzen-dorff@tu-clausthal.de

Abstract

Keywords: Care, Augmented Reality, Head Mounted Devices, Head Gestures

In this paper, we describe a study investigating augmented reality (AR) to support caregivers. We implemented a system called Care Lenses that supports various care tasks on AR head-mounted devices. For its application, one question was how caregivers could interact with the system while providing care, that is, while using one or both hands for care tasks. Therefore, we compared two mechanisms to interact with the CareLenses (handheld touch similar to touchpads and touchscreens and head gestures). We found that certain head gestures were difficult to apply in practice, but that except from this head gesture support was as usable and useful as handheld touch interaction, although the study participants were much more familiar with the handheld touch control. We conclude that head gestures can be a good means to enable AR support in care, and we provide design considerations to make them more applicable in practice.

1 Introduction: AR in (Home) Care

Care work is both important and a challenge in western societies, in which the population grows older. This results in a situation in which the number of people to receive care grows (Schorch, Wan, Randall and Wulf, 2016) and in which there is a shortage of care workers (Bratteteig and Wagner, 2013). As a result, caregivers need to provide care in less and less time. Using IT in care to disburden and support care workers has therefore been discussed widely. In this paper we describe a system called “Care Lenses”, which aims to support caregivers using augmented reality. This work is part of a larger research project that investigates the feasibility and potential benefits of using augmented reality support for care workers. The resulting system Care Lenses, a head mounted device (HMD) for AR support of care workers, is supposed to ease care work and enhance care quality. The project is based on an ethnographic study of care practices, from which we identified fields of support and matched them with the specific affordances of augmented reality (specifically HMDs), resulting in a set of features. Together with care workers we built working prototypes of these features and studied them in practice. In many care situations, care givers need one or both hands to interact with the patient or use assistive equipment. Therefore, we developed a mechanism that allows caregivers to control Care Lenses with head movements (“head gestures”), that is, without the need to use their hands for gestures or controls. This paper reports on a study in which we compared this mechanism with a built-in, touch-based controls of an HMD.

The main contribution of this paper is the presentation and evaluation of the head gesture interaction mechanism for HMDs in care practice. This is a novel concept that makes HMDs applicable in healthcare when both hands are needed for treatment, and our work shows that it works well in practice. In addition, the paper introduces Care Lenses and shows the potential benefit of the concept for the provision of care. To the knowledge of the authors, Care Lenses is unique among concepts applying HMDs in healthcare, as it supports a multitude of care tasks in practical care.

2 Related Work

2.1 Augmented Reality to Support Work

Using the words of Azuma (1997), in Augmented Reality (AR) “virtual objects superimpose[d] upon or composite[d] with the real world. Therefore, AR supplements reality, rather than completely replacing it.” By using AR devices, users enter what Milgram and Kishino (1994) have called “Mixed Reality” (MR), which merges digital and real worlds. AR has been found useful for the support of work in various domains and for various purposes, including remote (expert) support (Johnson, Gibson and Mutlu, 2015; Fakourfar et al., 2016), guidance (Büttner et al., 2017), remote cooperation (Datcu, Lukosch and Lukosch, 2016) and (remote) instruction or learning (Garrett, Jackson and Wilson, 2015; Preuveneers, 2015). The advantage of AR put forward in most of these examples is that it enables interaction with digital information and objects supporting work tasks while looking at the work scenery. AR can show information for tasks or annotations attached to items worked with (Fakourfar et al., 2016) as well as hints from a remote expert (Datcu et al., 2016). With AR, IT support therefore becomes an integral part of the work task and can be used easy and without frictions during work.

Augmented Reality can be implemented on different types of devices such as mobile phones, tablets or HMDs (often called “glasses” or “lenses”). While handhelds are commonly used and therefore offer better possibilities to access and use augmented reality, users need to use at least one hand (if not both hands) to hold and operate them. This makes it difficult to use AR in some of work areas, especially when tasks need both hands of a user (Johnson et al., 2015). An issue arising from HMDs, however, is how to interact with HMDs to tap from the potential of hands-free mixed reality. Features offered for this interaction often depend on the hardware available (cf. Schmalstieg and Hollerer, 2016). Common means of interaction available on current HMDs include swiping the frame of the HMD (e.g., Vuzix glasses), voice control, gestures and gaze (e.g., Microsoft HoloLens, Google Glass) or additional handheld devices like touchpads or clickers (e.g., Epson Moverio glasses and Microsoft HoloLens). With the exception of voice interaction, all of these mechanisms need at least one hand of the HMD user to be active in the interaction with the device. Recently, third party solutions for the usage of eye-tracking in AR have been developed and used for interaction with HMDs (e.g., Ku, Wu and Chen, 2017; Kytö et al., 2018), but there is no off-the-shelf solution for this modality except for single-gesture solutions such as blinking the eye on the Google Glass.

2.2 Head Gestures for Augmented Reality

A promising way of providing hands-free interaction is the usage of head movements as gestures to control AR and VR HMDs. By using built-in sensors of HMDs such as accelerometers and gyroscopes, distinct head movements can be interpreted and used as input commands. Head movements have been investigated as a mechanism to help people with disabilities control assistive technology (Jia, Hu, Lu and Yuan, 2007; Rudigkeit, Gebhard and Gräser, 2014). Besides this, research looks at head movements as an easy to use and precise way of pointing towards objects in mixed reality (Kytö et al., 2018), for authentication on HMDs (Yi et al., 2016), for the tracking of moving objects (Esteves et al., 2017), for mirroring and explicating emotions (Terven, Raducanu, Meza-de-Luna and Salas, 2016) and for interacting with AR HMDs (Starner, 2013; Yi et al., 2016).

Examples for gestures investigated for the interaction with virtual and augmented reality devices include nodding and shaking the head, turning it to the sides, looking up and down (and holding the head for a while after moving), tilting the head to the side, leaning forward or backward (including the upper part of the body), lines and geometric shapes (Ruban and Wood, 2016; Terven et al., 2016; Yi et al., 2016; Sharma, Ahmetovic, Jeni and Kitani, 2018). For the detection of head gestures, either motion sensors such as accelerometers and gyroscopes (Starner, 2013) or video analysis (e.g., Sharma et al., 2018) can be used. Using motion sensors has been shown to provide a low-cost but feasible method of detecting and discriminating gestures (Yi et al., 2016). Advantages of head gesture are that they are well known by many potential users of a system (thus easy to learn and use): they are used in everyday communication to convey meaning and can therefore be used intuitively by humans (Yi et al., 2016; Sharma et al., 2018). In addition, head gestures can be used easily and with good precision (Plaumann et al., 2015), and they provide a subtle means of interacting with a system (rather than e.g. voice control). Most important, head gestures can implement real "hands free" interaction with AR.

2.3 AR support for Care

Augmented Reality is becoming more and more interesting for research and development of support for work in healthcare. This, using the words of Siebert et al. (2017), is mainly because the potential of "freeing users' hands and allowing them to 'see the scene through the screen'" (see also Kobayashi et al., 2018). This shows the two main benefits for AR HMDs in care: First, they provide information during the care process rather than looking information up in the documentation and interrupting the care process for this, and (second) they leave the hands of caregivers available for the provision of care.

The majority of head-mounted AR (and VR) applications in care contexts deal with education (Garrett et al., 2015; e.g., Azimi et al., 2018; Kobayashi et al., 2018). Zhu et al. (2014) provide an overview of how AR and VR can support healthcare education, and Kobayashi et al. (2018) show that the areas of AR training are growing. For the usage of AR in medical training, Azimi and colleagues [2] show how AR based training enabled medical professionals to increase their "time-on-task" as well as their confidence in what they did. AR can also be used to provide remote (expert) guidance in care. Besides others, Mather et al. (2017) shows how their Helping Hands system can support caregivers in practice.

For the support of care tasks by AR there is only little work available. Among the applications of AR supporting care directly, Aldaz et al. (2015) present the SnapCat system, which uses Google Glass devices to make pictures of patients' wounds to support the documentation part of wound management in care. They argue that without their system, photo documentation of wounds needs at least two people, who position the patient to make the wound visible, hold a ruler to document the size of the wound, and take a picture. They find a preference of nurses for using their system over traditional means of wound management, and they attribute this mainly to hands-free documentation. Siebert et al. (2017) found that physicians using HMDs to guide defibrillation and other tasks in resuscitation simulations are as fast as colleagues using normal support but adhere much better to standard procedures and make less errors.

2.4 Open Issues and Research Question

The state of the art presented shows the potential of AR HMDs in care as well as the need for this support. However, there are only a few studies available for this support, and these are about specific features as shown above. In addition, little is known from these studies how caregivers are supposed to work with AR support while physically (with their hands, moving the body) interacting with patients. Therefore, questions remain such as how to provide caregivers with interactive support while providing care and how they can interact with HMDs while providing care to patients.

Our study was run to answer these questions. In particular, our work was directed by the following research question: How can head gestures facilitate the use of AR HMD support for care? To answer this question, we compared head gestures to the well-known interaction with the touch pad attached to the Epson glasses. In what follows, we describe the methodology of the study.

3 The CareLenses

3.1 The Concept of CareLenses

The CareLenses support care processes and enhance care quality. AR was used to overcome the obstacle that caregivers cannot use mobile devices or other material while they provide care as they need their hands for the provision of care and for hygiene reasons. The support provided is based on field work and co-design with caregivers. During the field work, different researchers conducted an ethnographic study with a total of ten patients, for whom we observed days or full shifts of care provided. In addition, we conducted 24 interviews with caregivers, care managers and relatives of patients. This informed the design of the CareLenses. First, we found that care workers' practices differ individually as well as from one organization or department to the other. We also found that care needs to follow guidelines to guarantee quality of care and well-being of the patient. In pain management, for example, patients are asked to provide an estimate of their pain level, and caregivers need to make sure the patient is conscious enough to provide this estimation. Therefore, support for care tasks cannot be provided in a strict step-by-step way but needs to allow caregivers to navigate flexibly through steps of their tasks while adhering to guidelines. Second, we saw that care is very personal and needs physical interaction with patients. These tasks need one or both hands of the caregiver to provide care, position the patient or use auxiliaries needed for the specific care task (see also Aldaz et al., 2015). This is a crucial constraint for the CareLenses, as support needs to be provided *while* care is provided, which means while the caregiver uses both hands. As an alternative, another caregiver could provide the support (from expertise or by reading from guidelines), but given shortage of personnel in care, such a helper is rarely available. From this, together with caregivers and experts, we derived a large set of support options to be provided by CareLenses, which includes care workflows (e.g., pain and wound management), documentation, ordering of assistive equipment and many others.

Figure 1. The Epson Moverio BT-300 HMD used for the Care Lenses (right: the handheld touch controller). Own image.

The CareLenses prototypes are implemented on an Epson Moverio BT-300 device (see Figure 1). This choice was made for multiple reasons. First, we wanted to use a binocular HMD, as monocular HMDs may lead to additional cognitive and physical demands (e.g., Matthies, Haescher, Alm and Urban, 2015). Second, many advanced HMDs like the Microsoft HoloLens or the Meta 2 glasses are large and look like helmets rather than supportive glasses. Caregivers told us this might interfere personal relations with their often vulnerable patients. Third, the Epson glasses use a handheld touch device attached to the glasses. As this interaction is known to users from their laptops and mobile phones, we expected the device to provide less of a burden in our initial design cycles than devices that would work with other controls for AR. In focus groups, however, the latter was also found to be problematic in cases in which caregivers needed both hands for care, which was the trigger for the work described here.

The CareLenses work in three phases: initiation, provision of support and documentation. Initiation is performed by either selecting a support feature from a menu or by context recognition. For the latter, the CareLenses recognize markers placed in the patient room or objects such as assistive equipment. From this, they deduct usage contexts. For example, if a pain scale is recognized, the lenses offer support for pain management or ordering pain scales (see Figure 2 below). For the provision of support, the CareLenses provide information for a task, step-by-step instructions for care tasks or access to organizational features. This support can be accessed and controlled by the caregiver during care. The documentation of tasks includes their completion and entering values (recognized by the CareLenses).

3.2 Sample Workflow used in the Study: Pain Management

The study described here used pain management workflow support as an example from the set of support available in CareLenses. We chose a medium complexity workflow as we did not want to create a burden by providing care givers with new technology *and* a workflow of high complexity. Pain management is an important task that we found to be error-prone (pain management) in practice.

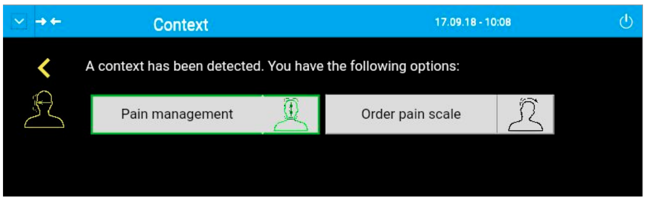


Figure 2. Starting the workflow by switching to the right button and selecting it (the black background of the screenshot is transparent when used on the CareLenses).

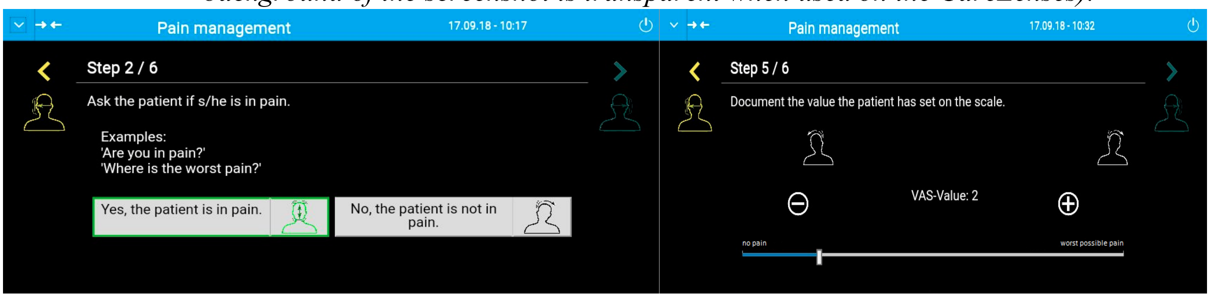


Figure 3. Left: Guidance for questions to use in pain management. Caregivers can approve or cancel the workflow. Right: Entering the pain level by tilting the head (icons right/left).

The pain management workflow includes seven steps (see Table 1). It starts as shown in Figure 2 with the pain management context detected by the e.g. the recognition of a pain scale. It then guides the caregiver through a process of asking patients whether they are in pain to having them assess their pain level (Figure 3 left) and documenting this level (Figure 3 right). As shown in Figure 3, pre-selected buttons allow caregivers to approve the conduction of steps and to proceed. After the assessment of patients' pain level, the caregiver enters it into the CareLenses by tilting her head left or right to set the respective value (Figure 3 right). To make the interaction more demanding and to simulate a mistake that happens in practice, we included a loop into this workflow. After entering the value for the pain selected by the patient into the dialogue at step 5 and reaching step 6, the patient told the caregiver he wanted to make a correction to the pain level. After that correction the caregiver had to go back to step 5 (called step 5_2 in Table 1), change the value and got to step 6 again (“End” in Table 1).

*Table 1: Steps of the workflow. Steps marked * include the tilt gesture shown in Table 2.*

Step	Description
Start	Starting the workflow (button pre-selected)
1	Approving that the patient can act for herself
2	Suggestion of questions for approaching patients
3	Handing out the pain scale
4	Receiving the pain scale
5/5_2*	Entering the pain level selected by the patient
6/End	Results of the pain management process

3.3 Two Interaction Concepts to support Care Work with the Care Lenses

Our first interaction mechanism used the *handheld touch* device that comes with the Epson Moverio device. The familiarity with this type of interaction provides a good baseline for other mechanisms to be tested with CareLenses. Despite this advantage, we assumed that the handheld was likely to create problems in practice, in which both hands are needed for the care task.

Being aware of the support needs and constraints in care, we wanted to provide a *real hands-free* interaction mechanism that allowed caregivers to use their hands permanently for care tasks. Modalities considered for this included eye tracking, head gestures and voice control. Among these, we discarded eye tracking, as at the time of our study there were only initial third party (add-on) devices for the control of AR HMDs with little research insights available. We then decided for *head gestures* (and thus against voice control) in order not to disturb the relationship between the caregiver and the (often-vulnerable) patient by having the caregiver speak commands into the HMD, which was also mentioned as a concern by caregivers in initial workshops. In contrast to voice commands, we considered gestures to be subtler with regard to disturbing interaction. In addition, we wanted to avoid accidental activation of commands during communication with the patient (cf. Yi et al., 2016). It should be noted that this was a decision for our study rather than a general assessment of the applicability of these mechanisms in care.

Gestures were detected by using the inertial sensors (accelerometer and gyroscope) of the glasses (see Yi et al., 2016 for a similar approach) to detect the movement direction and speed along the coordinate axes using the corresponding rotation rates. To avoid false detection by simply moving the head, we used pre-set thresholds for head rotation and changes of directions to detect gestures. These were set after a pre-test, in which users tried different movement speeds to find the most appealing configuration as in Yi et al. (2016). For example, we recognized nodding by detecting quick up and down movements of the head with at least two changes of directions (up, down, up or down, up, down).

For the set of gestures to be used, we wanted our gestures to be recognizable for the HMDs we used, and we wanted them to be as unobtrusive and natural (as they would be used in front of patients) and as intuitive as possible to lower the burden of using them (see Aldaz et al., 2015; Terven et al., 2016; Yi et al., 2016 for these requirements). Gesture candidates were taken from related work, considering results on recognizing, discriminating and using such gestures.

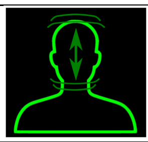
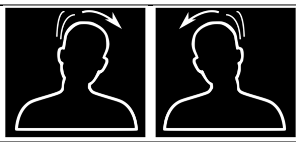
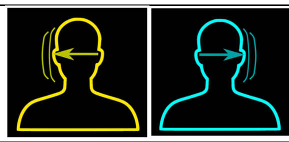
Gesture				
Description	Nodding	Tilting to the side	Shaking head	Turning to the side
Usage	Selecting, pushing button (approving)	Switching buttons / controls, setting values / scales	Cancelling, back to main	Back / forward in a workflow

Table 2. Gesture set for the head gestures of CareLenses. Icons taken from CareLenses UI.

From the workflow support to be used on CareLenses, we derived the commands to be covered by head gestures. These include starting a workflow and continuing to the next step (*approving*), going *back and forth* between steps, *switching* between buttons (for choices), *setting* values and *cancelling* a workflow. As the two most natural, distinct and easy to detect (Ruban and Wood, 2016) head gestures we selected nodding and shaking the head, which are intuitively associated with *approval* and disapproval or *cancelling* (Yi et al., 2016). We kept this association by using nodding for selecting and pushing buttons (e.g., acknowledging a step as in Figure 3) and head shaking for cancelling actions and returning to the home screen (see also Morency, Sidner, Lee and Darrell, 2007). For the remaining features such as going forward and backward between the steps in the workflows (depicted by the arrows in Figure 3) as well as selecting buttons or setting values (switching buttons in Figure 2, setting values in Figure 3), we selected turning to the side and tilting (see section 2.2), as they matched best to going back and forth in a workflow (turning the head, see also Jia et al. (2007)) and switching values (tilting the head, see also Crossan et al. (2009)). We noted that tilting and turning the head takes more time than nodding and shaking the head. We then adapted the temporal scale and thresholds for detection (Sharma et al., 2018).

We pre-tested the resulting gesture set (see Table 2) and found that the gestures could be executed and distinguished well by all participants of the pre-test. This also eliminated other gestures: for scrolling pages or moving up and down, we initially used ‘head up’ and ‘head down’ movements. In the pre-tests users had a hard time to differentiate these movements from nodding, and so we dismissed this gesture (see Sharma et al., 2018 for a similar observation). Rather than that, we avoided vertical menu structures.

4 The Study

4.1 Measurements Applied

To make sure we could compare the two mechanisms, we counted *user and mechanism-based errors*. We did this by looking at a combined video stream of the participant interacting with the patient and the corresponding screen shown by the CareLenses (Figure 4). Regarding user errors, we counted the number of times participants did not know how to proceed (e.g., when they did not know which gesture to apply in a specific situation) and their duration. For mechanism-based errors, we counted the times a participant acted correctly but the lenses did not react correctly (e.g., nodding without the mechanism recognizing it), asking the researchers what to do (e.g., how to proceed).

To shed light on the *effort to be taken in order to operate the CareLenses*, we applied the well-known (raw) TLX questionnaire (Hart and Staveland, 1988) and asked people to fill it in after each of the four tasks of the study. The TLX uses a scale from 0 to 20 to investigate work load dimensions of a task such as mental, physical and temporal demands as well as perceptions of performance, effort and frustration. It has been found to provide a good and valid means to access and compare task load induced by tool support for human work in healthcare and many other domains (Hart and Staveland, 1988; Hart, 2006). As usual in many studies, we used the raw TLX questionnaire without individually weights.

Besides errors and task load, answering our research question needs insights on the *task performance* resulting from the two mechanisms. We therefore took the times of task completion from the beginning (that is, when users started the task manually) to end of each task (that is, when they entered the final screen named “End” in Table 1). We also took the times spent by the participants on each step of the tasks, represented by the time each of the dialogues was visible to them. This way, we wanted to find out how well each of the mechanism supported specific steps.

As described above, *personal contact and empathy* are important aspects of providing care. Therefore, we analysed the videos taken during the studies and coded them with different categories to capture these dimensions. We applied continuous codes (during the task there was always one code active while the others were not selected) to depict whether the participants were directed towards the patient. We coded whether the body of the participant was turned towards the patient or not, and whether the participant was looking into the direction of the patient or not. We used this as an approximation of whether the participant paid attention to the patient or not (cf. Mehrabian, 2017). As codes were provided continuously, we counted the duration of each code. As a more direct measurement of personal contact, we coded whether the participant talked (or did not talk) to the patient. It should be noted, however, that in contrast to turning to patients the amount of time spent to talk to patients is highly individual, as some caregivers may use more words than others.

4.2 Course of the Study

The study was run with care providers from intensive care, home and elderly care. According to all participants, pain management is a relevant task for them. Table 3 shows the participants per provider.

Table 3. Care providers and participants in the study.

Provider	Participants
Elderly care ward	6
Intensive care shared apartments	6
Care laboratory, participants from different care providers	4
Intensive care stationary unit, participants from different providers	8

The age participants of the participants varied from younger than 25 to older than 50. 18 of them were female and six were male. The participants had on average 11.7 years of experience in care (StDev = 7.8). 15 of them dealt with pain management regularly, the other nine were familiar with it from nursing school. We looked for correlations between demographics of participants, but we did not find any medium or strong correlations. Thus, we assume that this did not have an influence on the results.

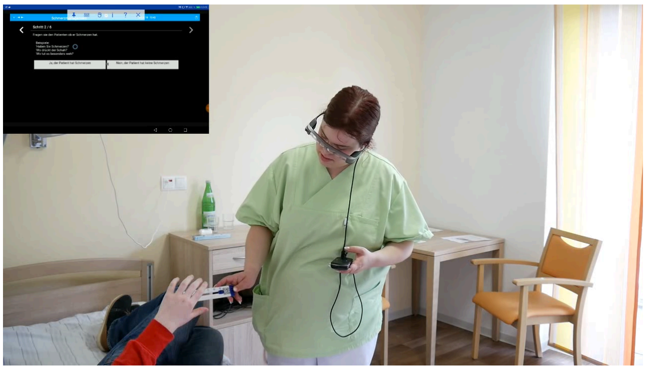


Figure 4. A participant in the study handing over the pain scale to the patient. The corresponding screen on the CareLenses is shown in the top left corner.

To provide external validity, the study was done in empty patient rooms. Participants were care workers on duty during the time of the study. They were allowed to take a break from regular work and participate in the study. They were in their care uniform and carried around items they use in daily work (see Figure 4). The main difference to real care tasks was that we did not work with real patients. One of the researchers took the role of the patient because the ethical approval we received from a national review board did not allow to test prototypes with real patients. While testing with real patients is preferable (and planned for later studies), this setting helped us to keep the behaviour of the patient constant (e.g., same answers to all participants when asked for pain levels etc.). For all tasks, we asked the caregivers to follow guidelines they received on the Care Lenses, even if they were used to a different procedure.

4.3 Study Design and Participants

We chose a counterbalanced within-subjects design to detect carry-over effects caused by using either a mechanism first, and to examine effects caused by familiarity with a task (that is, whether it was done the first or second time). Half of the participants started by using the handheld device followed by using gestures, and the other half started with gestures. The participants started with a brief tutorial for the respective interaction mechanism. Besides the tutorial and a brief explanation of tasks, the only instruction participants received was to complete tasks according to the guidance of Care Lenses. After each task was completed, the participants were provided with a TLX questionnaire.

It should be noticed that the comparison in this paper is clearly biased towards the usage of the handheld device. Interaction with handheld touch devices is well known and therefore this mechanism was likely to be more familiar to the participants. The head gestures, in contrast, were new to the participants and were therefore likely to be more difficult to use initially (cf. Plaumann et al., 2015). An extended training session for the gestures was not possible due to time restrictions in care organizations and may not have diminished this gap fully anyway. As a consequence, we did not expect the head gestures to outperform the handheld device but aimed for at least comparable performance.

5 Results

5.1 Data Analysis

For the analysis of the TLX questionnaire we used t-tests for paired samples, as data was normally distributed (Shapiro-Wilk test). For the usage and performance measurements, data was not normally distributed, and we used the Wilcoxon signed rank test for it. We checked the interrater reliability of our codes for performance and interaction by calculating interclass correlation coefficients (<.8 for all codes). We had to exclude some measurements on usage and performance for both conditions. This was done when participants did not follow the guidance on the screen due to misunderstandings (this happened especially for the first time they conducted the task) or when a mechanism error occurred that caused them to re-start the task when they had already completed it halfway or more (to ensure validity of comparisons). This way, the sample size went to 15 participants for these measures. However, the Wilcoxon signed rank test is robust against small sample sizes and our comparisons meet all requirements for alpha-values of 0.05 and 0.01 (see values reported below).

5.2 Usability

The data for the usability of the mechanisms shows that compared to the handheld condition, there were more user errors for the gesture condition ($z=-2.634$, $p<.001$, $T=8.0$, $n=15$). We also found more mechanism-based problems for gestures ($z=-3.313$, $p<.001$, $T=0$, $n=15$), longer times participants asked for help ($z=-2.04$, $p<.05$, $T=13.0$, $n=15$) and received help ($z=-2.09$, $p<.05$, $T=7.0$, $n=15$). Table 4 provides an overview of the corresponding values. There was only a significant difference in mechanism-based problems if gestures were used *before* handheld touch. This suggests that there were more difficulties with the gestures if the task was unknown.

Table 4. Results for error measures with significant differences (Wilcoxon signed rank test).

Category	Handheld	Gestures
Average number of user errors	1.87	4.6
Average number of mechanism problems	.07	6.4
t (user asking) in sec	1.8	7.8
t (help provided) in sec	2.0	15.5

The fact that the gesture mechanism was much more error prone than the handheld mechanism can be attributed to the novelty of the mechanism for the participants, which may have caused errors such as forgetting the correct gesture to process. In addition, we used a prototype which sometimes did not recognize gestures correctly. As Table 4 shows, this resulted in more time people spent with asking for help or receiving support (3.8 seconds on average for handheld, 23.3 seconds for gestures).

5.3 Task load

Regarding task load, we found significant differences in physical demand ($t(24)=-2.594$, $p<.05$), effort ($t(24)=-2.179$, $p<.05$) and frustration ($t(24)= -2.905$, $p<.05$). As Table 5 shows, all values reside in the lower levels of demand in the TLX scale (3 to 7). Looking deeper into these figures, we found that if handheld touch was used before gestures, there are no significant differences for any of the items. This means that if the task was known in advance, gesture interaction did not provide significant extra load.

The results for task load fit the results for usability in section 5.2. The higher load for gesture interaction may have been caused by more user and mechanism-based errors. Regarding the bias we assumed for the study (see 4.3) this was expected.

*Table 5. TLX scores. Results marked * differ significantly with a paired t-test (p<.05).*

TLX item	Handheld	Gestures
Mental demand*	1.56	3.44
Physical demand*	1.64	4.16
Temporal demand*	1.12	3.32
Effort *	.80	3.32
Overall performance*	1.64	3.6
Frustration	1.76	3.32

5.4 Performance during Tasks (original data)

All tasks were conducted correctly and finished successfully by the participants. However, we found significant differences in average overall and many times for steps. For the whole task, the participants needed on average 141.3 seconds with the handheld device and 183.3 seconds using gestures.

Table 6. Average times for tasks. All values differ significantly in a Wilcoxon signed rank test.

Step	t(handheld) in sec	t(gestures) in sec
5	17.3	38.6
Overall	141.3	183.3

As can be seen from Table 6, there is a significant difference in the time spent on step 5 of the task ($z=-2.442$, $p<.05$, $T=17$, $n=15$). Analysing the differences in task execution times, we found two important aspects. First, significant time differences appeared when participants had to use the tilting gesture (see Table 2) to set a value on the pain scale (step 5). When this gesture was not used, no significant differences were found in performance. In addition, when the tilting gesture was repeatedly used for the same task as in the loop between steps 5 and 6 (from step 5 to 6 to 5_2 as described in section 3.2), the difference in performance became much smaller, suggesting a learning effect for the gesture.

Second, the additional times spent for tasks in the gesture condition included the times for user and mechanism-based errors as well as for users asking and receiving help as shown in Table 4. Knowing these times is important to know that issues with using the mechanism and its proper functioning had an influence on task execution performance. However, these times also add on the overall times we compare and may therefore supersede other effects. In fact, we found that the majority of additional times spent were due to participants' brief exposure to the gesture mechanism (e.g., when they forgot one of the gestures) or to technical issues (e.g., when gestures were not recognized).

5.5 Performance after data cleaning

As we were interested in the performance apart from the issues with the tilting gesture, we cleaned the data by removing time needed to resolve unwanted (technical) issues. We re-coded the data, adding a code indicating that people spent time with usage issues rather than the task they were supposed to do. It was applied when participants started to focus on difficulties with the operation of the lenses and until the issues were resolved and participants re-entered task execution. We then re-calculated the times spent on the different tasks and steps in them and compared them again.

Selected performance data after data cleaning is shown in Table 7. Interestingly, we did not find any other significant differences in task performance pain after cleaning. For sequences in which the handheld condition was done before gestures, we found that the workflow was executed faster with gestures (handheld average 152.6 seconds, gestures average 127.2 seconds, $z=-2.032$, $p<.05$, $T=2.0$, $n=7$). Underpinning this, in linear regression analysis, we found significant effects of the sequence of conditions for total times spent on handheld and gesture conditions (handheld: $R^2=.297$, $F=9.864$, $\beta=-$

.575, $p<.01$; gestures: $R^2=.458$, $F=13.686$, $\beta=.703$, $p<.01$; residuals normally distributed in a Kolmogorov-Smirnov test). This suggests that for the cleaned data the differences in performance were mainly caused by learning effects when the task is performed twice rather than the mechanism used.

*Table 7. Left: Overall and step performances. Right: Personal interaction for different sequences of handheld and gesture. Values marked * differ significantly in a Wilcoxon signed rank test (p<.05).*

Step	t(handheld)	t(gestures)
All tasks		
Overall	135.9	147.6
Handheld before gestures		
Overall*	152.6	127.2

Interaction	t(handheld) in sec	t(gestures) in sec
Handheld before gestures		
Head/body tow. Patient*	67.2	21.3
Not talking to patient*	108.2	71
Gestures before handheld		
Head/body tow. Patient*	18.5	31.3
Not talking to patient*	76.1	123.4

5.6 Patient Interaction

For all participants and tasks, we did not find any significant differences in the interaction with patients. Looking at the different sequences of conditions provides better insights: If handheld was used before gestures, we found that for handheld the participants turned body and head much longer to patients (handheld average 67.2 seconds, gestures average 21.3, $z=-2.197$, $p<.05$, $T=1.0$, $n=7$) but spent longer times of not talking to patients (handheld average 108.2 seconds, gestures average 71 seconds, z=.2.201, $p<.05$, $T=1.0$, $n=7$). Table 7 shows this. The findings were reverse if the sequence was the other way round: Here, we found that participants using gestures looked longer at patients (handheld average 18.5 seconds, gestures average 31.3, $z=-2.1$, $p<.05$, $T=3.0$, $n=8$) but also spent longer times of not talking to patients (handheld average 76.1 seconds, gestures average 123.4 seconds, $z=-2.521$, $p<.05$, $T=0$, $n=8$).

Looking at this data it seems that if one mechanism was used first in the study, this resulted in more attention on the patient (head and body towards the patient) but less verbal interaction (longer times not *talking* to the patient). This could be an effect created by the sequence rather than by the mechanisms: once participants were more familiar with the task and its support on the CareLenses, they may have felt more comfortable when they used the CareLenses for the second time for this workflow and therefore they spent more attention to the patient. In the same way, the second time of task conduction may have resulted in a more routine and less explicit way to do the task, causing people to talk less. This is backed by regression analysis, in which we found a considerable effect on the attention to the patient for the handheld condition ($R^2=.231$, $F=7.325$, $\beta=-.518$, $p<.05$) and on not talking to the patient for both conditions (handheld: $R^2=.312$, $F=10.52$, $\beta=-.587$, $p<.01$; gestures: $R^2=.329$, $F=8.369$, $\beta=.612$, $p<.05$). All residuals were equally distributed in a Kolmogorov-Smirnov test. This suggests that the familiarity with the task was decisive for differences in the interaction with the patient rather than the mechanism.

5.7 Summary of the Results

Our results show that the gesture mechanism was much more error-prone than the handheld mechanism, which also resulted better TLX scores for the handheld mechanism. However, this was likely due to the prototype status of the mechanism, the tilting gesture that caused problems for some participants and the difference in familiarity between the two mechanisms. This interpretation is backed up by the fact

6 Discussion

6.1 Hands-free support by Head Gestures: Does it work?

In our research question (see 2.5), we asked to what extent head gestures could make hands-free AR support applicable in care practice. As an answer to this question, we can state that despite the bias for the more familiar touchpad interaction our head gestures often worked equally well. Gestures could be used intuitively and recognized reliably, which is in line with related work from other domains (Garrett et al., 2015; Yi et al., 2016; Azimi et al., 2018; Sharma et al., 2018). In other words, we compared a familiar and a novel interaction mechanism, and we found that the novel mechanism was not much worse in the worst case and better in the best case. We therefore conclude that head gestures are a promising way to provide such hands-free AR support in care. However, our findings also show that there are still issues to be dealt with. Among these, we found that some head gestures create difficulties for users, and that the gesture mechanism created more task load than the handheld. Therefore, there is still work to be done to use this interaction mechanism to provide hands-free AR interaction in care. In what follows, we will discuss our results.

6.2 The impact of familiarity with procedures and interaction mechanisms

In the analysis we often found that seeming advantages of the handheld mechanism observed such as faster execution for certain steps or TLX values handheld touch was used before gestures. In this condition, the task was already known to the participants from the handheld condition, and it seems that this made using the gestures easier for them. On the other hand, if both the workflow and the (gesture) interaction was unknown to participants, performance goes down. We conclude from this that with more familiarity of the workflow assistance in the CareLenses, gesture interaction can provide equal (or even better means) of interacting with AR assistance in care. The other way around, if participants knew neither the exact procedure they were to perform nor the modality of interaction, this was likely to amplify perceived and measurable difficulties with the interaction mechanism. In practice, this effect will most certainly diminish over time, as our findings on performance and other factors (cases in which gestures were used after the handheld) show. We therefore conclude that the head gestures can be a good means to control AR support for care tasks after a short period of getting familiar with it (as it is the case for most interaction mechanisms). However, some challenges remain and will be discussed below.

6.3 Gesture support is only as good as its worst gesture: Tilting Gesture

The number of errors and the differences in TLX scores (though residing on good to acceptable levels) show that the gesture mechanism was more difficult to use than the handheld mechanism. One big challenge our participants were faced with was the tilting gestures to switch buttons or set values. Whenever it was used (and only in these cases), the performance of the respective step was significantly worse. After data cleaning, which removed sequences in which the participants spend time with asking for and receiving help by the researchers, this difference was no more present to this extent. This suggests that tilting the head is a gesture that is less intuitive or much more difficult to perform correctly, thus lowering the usability of the mechanism and likely increasing task load. We conclude from this finding that gestures can be a proper mechanism for hands-free interaction with AR in care as long as we provide users with an easy to operate and semantically intuitive gesture set. While we will spend further work on making the tilting gesture easier for users, it seems likely that replacing is the better way to proceed. However, there is no semantically fitting head gesture left in the body of head gestures commonly used: Moving the head up and down was found to be too close to nodding, and head movements to *draw* gestures were found to be too difficult in practice (Yi et al., 2016). Solutions may therefore be found in mixed modalities or designing interfaces specifically for hands-free interaction.

6.4 Mixed modalities and designing for hands-free interaction

Our study suggests that simple head gestures such as nodding and turning the head to the side have the potential to provide good support for the support of care tasks. They worked smoothly in our study and (after we eliminated delays caused by user and mechanism errors) even outperformed handheld touch interaction in certain situations. As mentioned above, we also found limits to this support when it comes to more complex gestures like the tilting gesture we used for switching between options and values. As a solution for this, a mixed modalities approach that combines head gestures and speech input may be a good strategy. Such an approach could use the best of both mechanisms: It keeps subtle and easy-to-perform head gestures as the main control mechanism (e.g., choosing buttons) and adds speech input for speech commands such as entering numbers (e.g., the pain level selected by a patient as seen in Figure

3 right). In contrast to control commands, patients may understand why the caregiver says a certain number into a device if it matches the number they provided on the scale.

Using a mixed modalities approach, however, is not a complete solution for the issues we found. It does not solve selection problems such as switching to another branch of a workflow (e.g., stating the patient it not in pain as shown in Figure 3), and voice input may become clumsy when complex values need to be entered (including complex numbers and units). Besides mixed modalities, we may therefore strive to design specifically for hands-free interaction and control by head gestures. First, we could design the user interface for hands-free interaction, which is not the case if there are multiple buttons to choose from. Rather than this, head gesture control can work much better if there is only one choice to approve or to dismiss. For our case this would mean that the choice of the pain management workflow after context detection (Figure 2) may change: There would be only one button that offers to start the pain management workflow. Switching to the ordering workflow would then mean to dismiss this option by shaking the head, which would bring up a button to start the ordering process as a second option. This could be approved by nodding. This procedure would work for all steps discussed in this paper. To implement it, better context recognition to narrow down choices to a manageable number and a system of prioritizing options would be needed. For the latter, options with patient interactions could be preferred to make them accessible first. In tasks without this interaction, shaking the head once or twice may not be a huge problem and still lead to good task support quickly. Second, image and object recognition may enable the CareLenses to semi-automatically enter numbers displayed on devices or on the pain scale manipulated by the patient. This would only afford users to approve or dismiss the value read in automatically. Early work we are doing on the CareLenses shows that this is a promising approach.

6.5 Limitations

Our study was run in close-to realistic setting, but without real patients. As explained above, this was due to the ethical approval we received and the study setting, as we wanted the patient to behave in a certain way to control the study. This leaves out the patient’s perspective on and perception of AR devices in care. As mentioned above, we deliberately chose a medium-class HMD for the CareLenses, as it looks more natural than many high-class HMDs. However, how patients react to the CareLenses and how the react to our gesture mechanism remain questions to be answered. As mentioned in this paper, the work presented here is part of a larger stream of work, and evaluations with real patients are currently being planned (as planned by the ethical approval we received).

Besides this, the study was done with a sample of caregivers that – despite being representative for the care sector – is small could be extended to add to the generalizability of the results presented here. In addition, the task used did not go beyond medium complexity. This was a deliberate choice to not include too much complexity and burden on the participants (a very complex process and a new mechanism), and by the time of writing this paper, studies with complex tasks are being prepared.

7 Conclusion

In this paper, we have presented and explored head gestures as a means for hands-free interaction with AR support for care tasks. The need for such support arises from the complexity and quality affordances in care work, and by the need to use one or both hands for the conduction of care tasks. Our study with two typical care tasks shows that our head gestures performed equally well in many situations when compared to handheld touch controls, which we used as a baseline. From this we conclude that head gestures can be a means for hands-free AR control in care work. We also showed and discussed shortcomings of our mechanisms and open issues for the usage of head gestures in care work. We then presented potential ways to deal with these issues while keeping head gestures as the main interaction mechanism of the CareLenses. Our further work is set to implement and evaluate these improvements.

References

Aldaz, G., L. A. Shluzas, D. Pickham, O. Eris, J. Sadler, S. Joshi and L. Leifer. (2015). "Hands-Free Image Capture, Data Tagging and Transfer Using Google Glass: A Pilot Study for Improved Wound Care Management." PLOS ONE, 10(4), e0121179.

Azimi, E., A. Winkler, E. Tucker, L. Qian, M. Sharma, J. Doswell, … P. Kazanzides. (2018). "Evaluation of Optical See-Through Head-Mounted Displays in Training for Critical Care and Trauma." In: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (Vol. 00, pp. 1–9).

Azuma, R. T. (1997). "A survey of augmented reality." Presence, 6(4), 355–385.

Bratteteig, T. and I. Wagner. (2013). "Moving healthcare to the home: The work to make homecare work." In: O. Bertelsen, L. Ciolfi, & M. A. Grasso (Eds.), ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus (pp. 143–162). Springer.

Büttner, S., H. Mucha, M. Funk, T. Kosch, M. Aehnelt, S. Robert and C. Röcker. (2017). "The Design Space of Augmented and Virtual Reality Applications for Assistive Environments in Manufacturing: A Visual Approach." In: Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments (pp. 433–440). Island of Rhodes, Greece: ACM.

Crossan, A., M. McGill, S. Brewster and R. Murray-Smith. (2009). "Head tilting for interaction in mobile contexts." In: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '09 (p. 1). Bonn, Germany: ACM Press.

Datcu, D., S. G. Lukosch and H. K. Lukosch. (2016). "Handheld Augmented Reality for Distributed Collaborative Crime Scene Investigation." In: Proceedings of the 19th International Conference on Supporting Group Work (pp. 267–276). Sanibel Island, Florida, USA: ACM.

Esteves, A., D. Verweij, L. Suraiya, R. Islam, Y. Lee and I. Oakley. (2017). "SmoothMoves: Smooth Pursuits Head Movements for Augmented Reality." In: Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (pp. 167–178). New York, NY, USA: ACM.

Fakourfar, O., K. Ta, R. Tang, S. Bateman and A. Tang. (2016). "Stabilized Annotations for Mobile Remote Assistance." In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 1548–1560). Santa Clara, California, USA: ACM.

Garrett, B. M., C. Jackson and B. Wilson. (2015). "Augmented reality m-learning to enhance nursing skills acquisition in the clinical skills laboratory." Interactive Technology and Smart Education, 12(4), 298–314.

Hart, S. G. (2006). "NASA-Task Load Index (NASA-TLX); 20 Years Later." Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50(9), 904–908.

Hart, S. G. and L. E. Staveland. (1988). "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." In: Human Mental Workload (pp. 139–183). Amsterdam: North Holland.

Jia, P., H. H. Hu, T. Lu and K. Yuan. (2007). "Head gesture recognition for hands-free control of an intelligent wheelchair." Industrial Robot: An International Journal, 34(1), 60–68.

Johnson, S., M. Gibson and B. Mutlu. (2015). "Handheld or Handsfree?: Remote Collaboration via Lightweight Head-Mounted Displays and Handheld Devices." In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (pp. 1825–1836). ACM.

Kobayashi, L., X. C. Zhang, S. A. Collins, N. Karim and D. L. Merck. (2018). "Exploratory Application of Augmented Reality/Mixed Reality Devices for Acute Care Procedure Training." Western Journal of Emergency Medicine, 19(1), 158–164.

Ku, P.-S., T.-Y. Wu and M. Y. Chen. (2017). "EyeExpression: Exploring the Use of Eye Expressions As Hands-free Input for Virtual and Augmented Reality Devices." In: Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (pp. 60:1–60:2). New York, NY, USA: ACM.

Kytö, M., B. Ens, T. Piumsomboon, G. A. Lee and M. Billinghurst. (2018). "Pinpointing: Precise Head- and Eye-Based Target Selection for Augmented Reality." In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (pp. 81:1–81:14). New York, NY, USA: ACM.

Mather, C., T. Barnett, V. Broucek, A. Saunders, D. Grattidge and W. Huang. (2017). "Helping Hands: Using Augmented Reality to Provide Remote Guidance to Health Professionals." Studies in Health

Technology and Informatics, 241, 57–62.
Matthies, D. J., M. Haescher, R. Alm and B. Urban. (2015). “Properties Of A Peripheral Head-Mounted Display (PHMD).” In: International Conference on Human-Computer Interaction (pp. 208–213). Springer.
Mehrabian, A. (2017). Nonverbal communication. Routledge.
Milgram, P. and F. Kishino. (1994). “A taxonomy of mixed reality visual displays.” IEICE TRANSACTIONS on Information and Systems, 77(12), 1321–1329.
Morency, L.-P., C. Sidner, C. Lee and T. Darrell. (2007). “Head gestures for perceptual interfaces: The role of context in improving recognition.” Artificial Intelligence, 171(8–9), 568–585.
Plaumann, K., J. Ehlers, F. Geiselhart, G. Yuras, A. Huckauf and E. Rukzio. (2015). “Better Than You Think: Head Gestures for Mid Air Input.” In: Human-Computer Interaction (pp. 526–533). Springer.
Preuveneers, D. (2015). “The GhostHands UX: telementoring with hands-on augmented reality instruction.” In: Workshop Proceedings of the 11th International Conference on Intelligent Environments (Vol. 19, p. 236). IOS Press.
Ruban, T. and J. Wood. (2016). “Recognizing Head Gestures for Head-mounted Displays.”
Rudigkeit, N., M. Gebhard and A. Gräser. (2014). “Towards a user-friendly AHRS-based human-machine interface for a semi-autonomous robot.” In: 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, Workshop on Assistive Robotics for Individuals with Disabilities: HRI Issues and Beyond.
Schmalstieg, D. and T. Hollerer. (2016). Augmented Reality: Principles and Practice (01 ed.). Boston: Addison Wesley.
Schorch, M., L. Wan, D. W. Randall and V. Wulf. (2016). “Designing for those who are overlooked: Insider perspectives on care practices and cooperative work of elderly informal caregivers.” In: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (pp. 787–799). ACM.
Sharma, M., D. Ahmetovic, L. A. Jeni and K. M. Kitani. (2018). “Recognizing Visual Signatures of Spontaneous Head Gestures.” In: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV) (pp. 400–408). IEEE.
Siebert, J. N., F. Ehrler, A. Gervaix, K. Haddad, L. Lacroix, P. Schrurs, … S. Manzano. (2017). “Adherence to AHA Guidelines When Adapted for Augmented Reality Glasses for Assisted Pediatric Cardiopulmonary Resuscitation: A Randomized Controlled Trial.” Journal of Medical Internet Research, 19(5).
Starner, T. (2013). “Project glass: An extension of the self.” IEEE Pervasive Computing, (2), 14–16.
Terven, J. R., B. Raducanu, M. E. Meza-de-Luna and J. Salas. (2016). “Head-gestures mirroring detection in dyadic social interactions with computer vision-based wearable devices.” Neurocomputing, 175, 866–876.
Yi, S., Z. Qin, E. Novak, Y. Yin and Q. Li. (2016). “Glassgesture: Exploring head gesture interface of smart glasses.” In: Computer Communications, IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on (pp. 1–9). IEEE.
Zhu, E., A. Hadadgar, I. Masiello and N. Zary. (2014). “Augmented reality in healthcare education: an integrative review.” PeerJ, 2, e469–e469.