Comparing Smartwatch Input Modalities for Older Users in a TV-Control Scenario

Jonas Roberts, Johannes Wölfl, Christoph Märkl, Patricia Böhm, Daniel Isemann

Chair of Media Informatics, University of Regensburg

Abstract

Older people have been hesitant to accept smartwatches as supporting devices in their daily routine. We present a study which used an environment familiar to many older users to examine differences between touch and voice input for smartwatches. As people get older they tend to spend considerable amounts of time watching television. In this familiar setup we examined to what extent touch and voice input differ, using a smartwatch as a television remote control. We developed two smartwatch applications, one for each input modality, with which users could operate a TV simulator. In pre-tests, we determined possible speech input commands and preferred smartwatch gestures. Testing with eight older adults we found potential for both input modalities with speech input taking significantly longer. Although a majority of participants reported speech to be "easier" (with the rest undecided), participants were ultimately divided in their preference of one system over the other.

1 Introduction

There is a documented lack of acceptance among older people towards devices such as smartwatches when they are used for notifying them on being late on their medicine taking or monitoring them during their walks. This is likely, because users perceive this as surveillance and fear the social isolation when they are communicating with a smart device instead of a real person (Angelini et al., 2013).

We investigated smartwatch interaction in a very common and familiar environment: watching television. As people get older, they tend to watch more television. Studies suggest, that Germans over 50 on average watch 299 minutes of television per day (Media Control, 2015).

Veröffentlicht durch die Gesellschaft für Informatik e. V. 2017 in
M. Burghardt, R. Wimmer, C. Wolff, C. Womser-Hacker (Hrsg.):
Mensch und Computer 2017 – Workshopband, 10.–13. September 2017, Regensburg.
 https://doi.org/10.18420/muc2017-ws17-0413

While touch input is more popular than speech input among elderly users, this preference changes when it comes to smaller displays like smartphones (Teixeira et al., 2012). When designing applications for even smaller displays it is therefore important to look closely at different input options in order to draw the right conclusions and arrive at a user friendly solution for older people.

We developed two smartwatch applications, one operated by touch and one operated by speech, to control and navigate a (simulated) television system. To review our setup and to identify potential problems, we conducted a pre-test with young adults. Subsequently, we conducted an experiment to examine the differences between touch and voice input in terms of usability and individual preferences.

2 Related Work

Our study focusses on older adult users. As populations are growing older, the cost of healthcare is increasing. It has been suggested that investigating possible solutions for improving home care and autonomy of senior citizens may help alleviate healthcare costs (Ehrler et al., 2014). Ehler and Lovis examined the advantages and disadvantages of using smartwatches at home to provide assistance with medicine taking. They conclude the fact that smartwatches are continuously worn on the person to be an advantage in case of accidents. On the other hand, they identify as disadvantages the lack of acceptance of smartwatches by older people and the need for larger UI elements on the small smartwatch displays. Addressing the latter problem, Angelini et al. (2013) designed a smart bracelet with simple iconic images as UI elements. They report that users found the smart bracelet UI useful but were also concerned because they perceived it as a surveillance device.

In order to overcome lacklustre acceptance of digital devices such as smartwatches Kriglstein & Wallner (2005) experimented with a stuffed animal low-fidelity prototype of an artifical companion. The companion was voice-controlled and connected to a TV set. It served to evaluate a user-centered design of an interactive assistance system for the elderly. Among the subjects, the interaction with the prototype and way of controlling the television found high acceptance.

Teixeira et al. (2012) developed a mobile application and a touch PC application enabling social media services and evaluated both with older users. They conclude that touch controls are preferred over spoken input. On mobile devices, however, touch input is less popular because of the small display and keyboard.

3 Methods

In our study two smartwatch apps were developed to compare touch and speech as input modalities for the elderly. Speech input commands and touch gestures were identified in interviews with older subjects and evaluated in pretests with 6 younger adults. The resulting

functional prototype (see Section 4) was evaluated in field trials at the subjects' homes with eight older participants (4 male, 4 female) aged between 57 and 73.

Each subject completed the following tasks in a within subjects design with both modalities: 1) turn on television 2) adjust the volume 3) switch to a specific channel 4) switch to the next channel 5) activate information 6) turn off television. During the test, we collected error rate, completion rate and completion time. After the evaluation of one modality, we asked 3 questions about the acceptance answered on a 5-point Likert scale (1 - “I strongly disagree” to 5 - “I strongly agree”): 1) Would you recommend others in your age group to use a smartwatch with voice input/touch input? 2) The use of a smartwatch with voice input/touch input could have positive effects on my everyday life. 3) Using a smartwatch via voice input/touch input could increase my productivity in everyday life. After that the test persons rated the usability of the used input modality with the System Usability Scale questionnaire (Sauro, 2011). Finally, we asked the subjects some concluding questions about the presented modalities concerning possible use cases in everyday life, personal preference, ease of use, and which method they would be more likely to use on a daily basis.

4 Technical Setup and TV Simulator

Our technical setup consists of three main elements: A TV screen, a TV simulator application running on a laptop and Smart Watch App (see Figure 1). The TV simulator is implemented as a Windows application. It simulates the linear TV programme for 10 selected german television channels playing a looping, pre-recorded afternoon show. The task data is logged by the software, along with participant metadata. On the TV screen instructions for each upcoming task and the simulated "TV programme" were shown. A HTTP API server allowed the smartwatch application to control the TV simulator wirelessly. A chatbot implementation interpreted and executed the speech input result from the smartwatch. The speech recognition component is based on the standard package for Android.

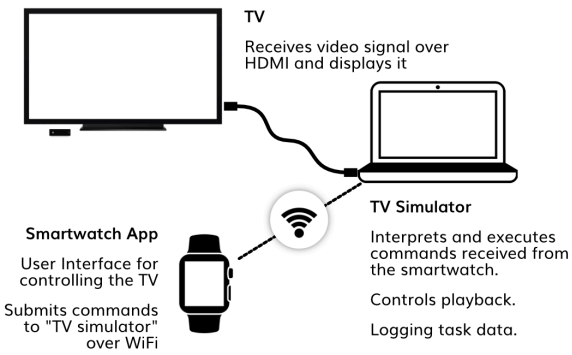

Figure 1: Technical setup

The two smartwatch applications were able to operate or "control" the TV Simulator running on a Laptop and thereby allowed the user to perform basic tasks such as turning the TV on and off, changing the volume and channels and displaying information about the currently selected channel on the TV. One version of the app was used to conduct our experiment with regular touch input (see Figure 2). A modified version was used for speech input and consisted of a single screen inviting the user to start the speech recognition process (see Figure 3).

Figure 2: Smartwatch touch application

Figure 3: Smartwatch speech application

5 Results

To examine efficiency of the two input modalities task times were collected. The mean completion times of several tasks (especially Task 1, 3, 5 and 6) indicate, that tasks were completed faster with touch interaction (see Table 1). To determine statistical significance of this tendencies we conducted a paired t-test on the average task times for all tasks (N = 48). To test significance for each task separately our sample size was too small (N = 8). We used a Log Transformation on the skewed data to ensure normal distribution. The Shapiro Wilk test confirmed that the both samples have a gaussian distribution. The paired t-test on mean completion times for *speech* (M = 12.24s, s.d 9.65s) and *touch* (M = 9.40s, s.d. 7.28s) revealed a statistical difference (t-value: 2.156, p = 0.036). The subjects in our study were significantly faster using the touch input.

The mean System Usability Scale (SUS) for touch input (M = 67.81, s.d. 22.2) and speech input (M = 67.19, s.d. 30.2) are more or less identical. According to a meta research summarizing the SUS results of 500 studies, a score of 68 is average or can be interpreted as a grade "C" (Sauro 2011). Our mean scores indicate that older subjects perceived the usability of both input modalities as "satisfying".

To assess the acceptance of touch and speech input for smartwatches three items concerning the recommendation to others, positive effects in everyday life and increased productivity in everyday life, were asked. The answers were given on a 5-point Likert scale. The resulting mean scores (see Table 2) for all three items and both conditions are all above 3.0. The

average for all three items is 3.50 (s.d. 1.285) for speech and 3.46 (s.d. 0.884) for touch. These scores indicate a small tendency to accept both input modalities.

Task	*Modality*	*Mean completion time*	*Median completion time*
(1) turn on television	*speech*	12.4 (s.d. 5.3)	10.8
	touch	**8.7** (s.d. 6.1)	7.1
(2) adjust the volume	*speech*	15.9 (s.d. 8.7)	14.1
	touch	17.4 (s.d. 8.8)	20.9
(3) switch to a specific channel	*speech*	**12.1** (s.d. 12.7)	7.5
	touch	**5.0** (s.d. 1.4)	4.9
(4) switch to the next channel	*speech*	10.0 (s.d. 3.0)	10.6
	touch	**9.6** (s.d. 3.3)	8.8
(5) activate information	*speech*	19.4 (s.d. 14.3)	13.3
	touch	**14.9** (s.d. 8.9)	11.4
(6) turn off television	*speech*	18.0 (s.d. 8.3)	17.5
	touch	**12.6** (s.d. 5.9)	12.5

Table 1: Task completion times

Asked about their preferred input method, four out of eight participants preferred touch input. Six participants found speech input to be easier to use, while one could not decide and one did not find any input method to be easy. For everyday use, four users would choose touch input over speech, two preferred speech input and one person would not use any of the input methods.

While two participants could not imagine themselves using a smartwatch, four thought it could be used to control home appliances as well as to check the current status of such devices. Three users suggested to control the roller shutters remotely for convenience.

Question	*Mod.*	*N*	*Mean*	*Median*
(1) recommend modality	*speech*	8	3.13 (s.d. 1.55)	4.00
	touch	8	**3.63** (s.d. .52)	4.00
(2) positive effects	*speech*	8	3.50 (s.d. .76)	4.00
	touch	8	3.50 (s.d. 1.07)	4.00
(3) increase productivity	*speech*	8	**3.88** (s.d. 1.46)	4.50
	touch	8	3.25 (s.d. 1.04)	3.00

Table 2: Results for questions about acceptance

6 Discussion

We compared touch and speech as smartwatch input modalities for the elderly and identified issues as well as potentials for both interaction forms. The statistical analysis of the collected task completion times suggests, that speech input takes significantly longer than touch input on smartwatches due to high latencies in the recognition. Touch input also posed some problems. Several subjects had difficulties with switching visual focus from the smartwatch to the rather distant TV and had to adjust their glasses. The comparable mean scores for the acceptance (3.50 for speech and 3.46 for touch) indicate that older users see potential in both modalities. Although speech was considered easier, in everyday use, touch input was preferred. Six out of eight participants could imagine to control other devices like home appliance or roller shutters with a smartwatch.

Our results could be used for further research on the suggested usage scenarios and input modalities. We recommend to use *Wizard of Oz* experiments to mitigate the off-putting effects of speech technologies latency. We refrained from using a *Wizard of Oz* setup to evaluate the practical rather than the theoretical usability and acceptance of the current state of the technology and how it might be used by elderly in everyday life.

7 Literature

Angelini, L., Caon, M., Carrino, S., Bergeron, L., Nyffeler, N., Jean-Mairet, M., & Mugellini, E. (2013). Designing a desirable smart bracelet for older adults. *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication - UbiComp '13* Adjunct, 425–434. https://doi.org/10.1145/2494091.2495974

Ehrler, F., & Lovis, C. (2014). Supporting Elderly Homecare with Smartwatches: Advantages and Drawbacks. *Studies in Health Technology and Informatics, 205*, 667–671. https://doi.org/10.3233/978-1-61499-432-9-667

Kriglstein, S., & Wallner, G. (2005). HOMIE: an artificial companion for elderly people. *CHI '05 Extended Abstracts on Human Factors in Computing Systems - CHI '05*, 2094. https://doi.org/10.1145/1056808.1057106

Media Control. (2015). Durchschnittliche tägliche Fernsehdauer in Deutschland nach Altersgruppen in den Jahren 2014 und 2015 (in Minuten). *Statista - Das Statistik-Portal*. https://de.statista.com/statistik/daten/studie/222521/umfrage/taegliche-fernsehdauer-nach-altersgruppen-quartalszahlen/

Sauro, J. (2011). Measuring Usability with the System Usability Scale (SUS). Retrieved from https://measuringu.com/sus/

Teixeira, V. D., Galinho Pires, C., Pinto, F. M., Freitas, J., Sales Dias, M., & Mendes Rodrigues, E. (2012). Towards Elderly Social Integration using a Multimodal Human-computer Interface. *2nd International Living Usability Lab Workshop on AAL Latest Solutions, Trends and Applications - AAL 2012*. Vilamoura, Algarve, Portugal, 3–13.