Usability and User Experience, Vol. 39, 2022, 194–200
https://doi.org/10.54941/ahfe1001708

Voice Interaction Design of Intelligent Wearable Device in Elderly Exercise and Health Scene

Li Yijun and Zhang Hui

School of Design, Hunan University, Changsha, Hunan, China

ABSTRACT

Health monitoring is a significant application scenario of intelligent wearable devices. In the era of an aging society and intelligent elderly care, the weak physical function of the elderly will drive their demand for intelligent wearable devices and become an important user group in this field. Related products will carry out differentiated designs according to the characteristics and demands of elderly user groups. Under the vertical segmentation of scenes, the user experience will continue to upgrade, and the continuous innovation of interaction methods is also accelerating. Especially with the development of technology, AI improves interaction ability and data value, and the requirements for voice interaction are improved. Voice is the natural interaction channel of intelligent wearable devices and the natural interaction mode of the elderly. However, at present, less research focuses on the voice interaction of smartwatches, especially on the sound details in different scenes. Based on the elderly exercise and health scene, this study explores the voice interaction design of smartwatches to put forward the corresponding voice interaction model and user experience design suggestions, to promote reference in the field of voice interaction of intelligent wearable devices.

Keywords: Intelligent wearable devices, Smart watches, Voice interaction, Aging design, User centered design

INTRODUCTION

With the development of sensor and AI technology, the application scenarios of intelligent wearable devices are becoming more and more mature. Digital exercise and health functions lead the upgrading direction of smartwatches and have become the main selling point. At the same time, with the rising trend of China's aging population and the advent of the big data era, the health care of the elderly is gradually transferred to be more intelligent with the further developed smart elderly-care industry (Qin 2021). As typical products of intelligent wearable devices that can provide good health management services, smartwatches have the advantages of lightness, portability, and closeness. Smartwatches have a smaller volume, weaker screen input function, and higher requirements for sound interaction. AI and voice are combined to form intelligent assistants embedded into devices. It can actively extract practical information according to its database and continuously

upgrade through self-learning to provide users with a more convenient and straightforward system. It can release users' hands, enabling users to realize convenient interaction and improve the experience in different scenes. Furthermore, in the case of physical function degradation such as vision loss, voice interaction has naturally become a more suitable interaction mode for the elderly (Wang 2019).

However, the user experience research of voice interaction pays less attention to the elderly. Although lots of research focuses on intelligent wearable devices (Esfahani and Ganji 2021, Zhou and Zhou 2021, Sarkar and Chakrabarti 2021), there still need more supplements on the related voice interaction design. Designers had not fully considered the difficulties faced by the elderly in voice interaction with different devices and scenes, and the user experience needs to be improved. From the perspectives of scene design (Sun and Wu 2017, Jiang 2020), user-centered-design, and emotional design, this paper analyzes the cognitive and psychological characteristics of the elderly. It explores the elderly's needs for exercise and voice interaction through field observation and semi-structured interview research. Then put forward the corresponding voice interaction design model and strategic suggestions to improve the user experience, especially in the exercise and health scene, which can provide some references for the aging design of voice interaction based on intelligent wearable devices.

BACKGROUND AND OVERVIEW

Overview of Smart Wearable Devices and Voice Interaction

Intelligent wearable devices comprehensively use various identification, sensing technology, cloud service, other interaction, and storage technologies to realize user interaction, live entertainment, human body monitoring, and other functions. According to the product form, the related industry can be divided into smart glasses, smartwatches, smart bracelets, smart helmets (Peng et al. 2017), etc. Smartwatches use the built-in intelligent system to connect the network to realize its function, synchronizing the phone, SMS, email, photos, and music in the mobile phone. The embedded sensors can also accurately track the changes in human indicators. With the maturity of sensor and AI technology, the leading smartwatches on the market have rich sports and health functions, which are mainly divided into two modules: (1) Digital health. Monitor the heart rate, ECG, blood pressure, blood oxygen, and other indicators, analyze the user's health status and guide the healthy life according to the intelligent algorithm; (2) Sports data analysis. Monitor the amount of exercise in swimming, running, mountaineering, and other activities, and analyze the indicators such as sports distance and calorie consumption according to the algorithm to assist users scientifically and effectively.

Voice interaction (Harris 2004) refers to the information transmission between the user and the system through natural language. The user and the system determine the design of the voice interface. A complete speech interaction includes three main parts: speech recognition, natural language processing, and speech synthesis, which continuously cycles to complete the

whole process of human-computer interaction. At present, the widely used voice interaction framework is mainly divided into four parts: awake, input, understand, and feedback. Awake means that the user wakes up the system with awake words or touch buttons; Input refers to the user's voice input through speaking; Understand means that the system understands the meaning of the voice input by the user for subsequent response; Feedback refers to that after decoding the user's language, the system makes corresponding instruction feedback, and conveys the results through voice or screen display.

Voice Interaction of Intelligent Wearable Devices

The voice control and interaction technology system of intelligent wearable devices is critical to liberating users' hands. Many mainstream smartwatches, such as Apple Watch, Huawei Watch, and Xiaomi Watch, have embedded intelligent voice assistant in the system, which can not only be used for listening to songs, making phone calls, navigating, setting the alarm clock, checking the weather and turning on sports mode, but also serve as the access port and operation panel of smart home, which can turn on the air purifier, turn off the fan, turn on the desk lamp, turn on the TV, etc. Multimodal interaction (Turk 2014) is involved in the practical application scenario of intelligent wearable devices. In addition to voice interaction, the whole human-computer interaction process will also cooperate with other interaction processes and feedback, such as biological information input, vibration feedback, etc. Taking the smartwatch as an example, with the help of the sensor, it can also trigger the wake-up process through the user's health indicators and other information or automatically input command information through the sensor. After decoding the input information, the system can give interactive feedback via voice, screen display, and vibration.

With the trend of an aging society and the development of intelligent elderly-care, the weak physical function of the elderly will stimulate their demand for intelligent wearable devices and become an important user group in this field. Therefore, it is vital to carry out differentiated product design according to the characteristics and demands of the elders, and the products will become more intelligent, emotional, and universal. Voice interaction has the following advantages for the elders: freeing hands and eyes, easing the pressure of action, using intuitive interaction, low learning cost, and high input efficiency. It also has the following disadvantages: increasing the burden of information reception, vulnerability to environmental interference, and fuzzy information receiving nodes. Using its advantages and optimizing its weaknesses can improve the user experience of voice interaction products. In each interaction stage, the cognitive differences of the elders will impact their experiences. Therefore, we need to fully consider the elders' characteristics and the merits and demerits of voice interaction in the design process.

RESEARCH PROGRESS

Analysis of The Elderly's Basic Characteristics

With the growth of age, people's bodies are inevitably aging, and the physiological, cognitive, and psychological characteristics in aging are also

Voice interaction phase	Physiological characteristics changes	Cognitive performance	Psychological performance	Design insights
Awake	Decline of nervous system function.	Difficult to remember the awake ways and words.	Distrust of smart devices and emotional tension.	Simple awake words, obvious button. Automatically monitor the situation and wake up.
Input	Cognitive decline.	Difficult to express clearly and fluently.	Feel frustrated about unfluently expression.	Sufficient response time. Gentle and patient guidance.
Understand	Cognitive decline.	Difficult to express clearly and fluently.	Tends to produce anxiety.	Gental comforts for waiting. Disclosure inquiry.
Feedback	Decline of hearing, memory and comprehension.	Poor understanding of feedback information.	Feel negative and resistant, and think that the product can not meet the requirements or the function is too complex.	Clear and accurate feedback. Repeat confirmation if necessary.

Figure 1: Cognitive load of voice interaction for the elderly and related design insights.

changing. As shown in Figure 1, the study summarizes the elderly's cognitive load of voice interactive (Wang and Yin 2021) by understanding their basic features to guide the design better and meet their needs.

Analysis of The Elderly's Physical Exercise

Relevant research (Yang et al. 2019) shows that among the primary physical exercise activities of the elderly in China, the top three are "walking," "dance (square dance, social dance, folk dance, etc.)," and "running." These activities are less limited by venues and facilities. They have lower requirements of complicated skills and professional sports equipment, so they are more suitable for the elderly in terms of physical quality and economic conditions. The elderly usually do jogging, Taijiquan, and aerobics from 6:00 to 7:30 in the morning. Evening exercise (about 2 hours after dinner) is the most common, including popular activities such as running, dancing, aerobics and so on, especially square dance. The elderly primarily exercises in parks and squares. They generally lack professional knowledge and guidance, so they often choose sports projects that do not need professional equipment and expertise. Therefore, parks and squares have become suitable venues for the majority of the elderly.

User Interview Research: Demand Analysis on Exercises and Voice Interaction of the Elderly

User interview is one of the most common qualitative research methods. Through field observation and semi-structured interviews of 12 elderly, the study explores the elderly's needs in exercise and health scenes to understand the main motivations and the pain points of voice interaction when using intelligent wearable devices. The interview contents mainly include the following topics: First, understand the basic information, daily life, and exercise of the elderly. Second, understand the elderly's acceptance and familiarity with voice interaction; Third, understand the elderly's relevant thoughts on the assisted activity of intelligent wearable devices. We sorted out the corresponding demand pain points (Figure 2).

In general, the elderly pursue simplicity and convenience in the functional needs of exercise and voice interaction. It is necessary to avoid complex operations and reduce the understanding cost of the elderly as much as possible

Aspects	Pain points and demands	Frequency
Functions of exercise and health	• The operation is simple, intelligent and automatic	10 people
	• Able to monitor physical health indicators (such as heart rate)	10 people
	• Be able to give relevant reminders during exercise	9 people
	• Be able to carry out exercise teaching guidance	8 people
Voice interaction	• The operation is simple and not complicated	12 people
	• Can understand accurately	10 people
	• Give clear and accurate feedback	10 people
	• Loud (audible)	9 people
	• Speak slowly (moderately)	8 people
	• Less screen operation	8 people
	• The screen can be seen clearly and the contents are large enough	7 people
	• The voice is kind, positive and friendly (sounds comfortable)	7 people
	• The expression is easy to understand	6 people

Figure 2: Frequency of the elderly's pain points and demands.

in the design progress. Another matter is to clearly and accurately understand the input information from the elderly and give clear, accurate feedback to hear, understand and react. It should avoid the screen operation as far as possible. If unavoidable, the screen should be more extensive, and the displayed context should be large enough to be seen. At the same time, in the field of exercise and the health of the elderly, it needs to meet the functional demands. It needs to consider the emotional orders and convey humanized care to make the elderly emotionally satisfied. In terms of voice and dialogue, it should be as kind, positive, and friendly as possible so that the elderly can feel cared for and try to eliminate their loneliness and sense of exclusion.

As suitable smartwatches, they should meet the exercise demands of the elderly for significant activities such as walking, running, square dancing, etc. It needs good functions such as monitoring and reminding, teaching guidance, recording, and feedback. When it comes to voice interaction, a good voice interface requires not only excellent natural voice understanding but also strategies to help users understand the range of actions and commands available in the interaction process, which needs to consider the following voice interaction design principles: Predict and confirm the user's intention; Give appropriate implicit prompt; Ask in turn to reduce the cognitive load of users; Progressive disclosure, providing rarely used options only after user behavior indicates that these commands are needed; Flexible and efficient system, to be convenient for voice interaction in exercise and health scene.

VOICE INTERACTION DESIGN MODEL

Through the analysis of the types of exercise activities and relevant needs of the elderly in the previous chapters, based on the exercise functions supported by the mainstream smartwatches on the market, the voice interaction motivation in the elderly exercise and health scene is divided into five categories: teaching guidance, health monitoring, exercise feedback, emergency reminder, and common reminder. Based on the previous research and analysis on the elderly's exercises conditions and voice interaction demands, according to the voice interaction motivation in the elderly exercise and health

			Voice interaction motivations in exercise and health scenes				
			Types of motivation				
			Teaching guidance	Health monitoring	Exercise feedback	Emergency reminder	Common reminder
Motivational behaviors			Professional teaching guidance on exercise behaviors	Monitor and remind physical indicators (heart rate, pressure, respiration…)	• Recording feedback • Encouraging feedback • Suggestions feedback • Stop Feedback	Emergency reminders of falls, injuries, physical discomfort and other emergencies	• Prompt for social information • Prompt for system notification
Interactive content example			"Please lift your left hand slowly~"	"The current heart rate is too high, please pay attention to rest!"	• "10000 steps!" • "Perfect action!" • "Let's take a break~" • "Please stop right now!"	Actively confirm the situation and call for help in time	• Vibration + voice • Vibration (can be set to off)
VOICE INTERACTION DESIGN	Role		Male / female. Positive, Kind, Professional, Reliable. (It is recommended to provide male and female voice gender in the actual product design, so that users can make personalized choices.)				
	VUI	sound tone	Higher	Moderate	1. Moderate 2. Higher 3. Moderate 4. Lower	Higher	Moderate
		speed	Slower	Moderate	Slower	Slower	Moderate
		volume	Louder	Louder	Louder	Louder	Moderate
		emotion	Cordial, Professional	Professional, Serious, Kind	1. Cordial & Kind, 2. Happy 3. Professional, 4. Serious	Professional, Serious Gentle, Kind	Cordial, Standard
		dialogue	Emotional, concise and clear, easy to understand, avoid complex and unfamiliar professional words, repeated confirmation of intention when necessary, and reserve enough understanding time and reaction time.				
	GUI		Necessary schematic icons, large texts & icons, clear visibility, strong contrast, high color saturation; If it is unnecessary, try to avoid screen display and operation, but at the critical feedback points, the feedback results can with graphical illustration; The pop-up interface prompt can be closed automatically.				
	Other feedbacks		Users can adjust the amplitude of vibration feedback. Vibration feedback can strengthen the reminder in the key action steps. If needed, flicker feedback also works (such as enhancing the vibration amplitude and flicker effect in case of stopping feedback or emergency and enhancing the reminder effect). Appropriate vibration feedback can help remind important messages (such as marked important phones, SMS, etc.).				

Figure 3: Voice interaction design model in elderly's exercise and health scene.

scene, focuses on the perception part of the user's perspective in the voice interaction process, divide the content of voice interaction design into several parts: role design (gender, personality) (Pradhan et al. 2018, Ji et al. 2019, Niu 2021), voice interface design (sound feature design, dialogue content design) (Alam and Riccardi 2014, Silnitskaya and Gusev 2013), visual interface design and other interactive feedback. Then propose the voice interaction design model of intelligent wearable devices in the elderly exercise and health scene (Figure 3).

DISCUSSION AND CONCLUSION

This paper focuses on the user experience needs of intelligent wearable devices in the elderly's exercise and health scene, studies smartwatches' aging voice interaction design, and puts forward the corresponding voice interaction design model and user experience design suggestions. Due to the influence of sample size, user cultural differences, language differences, and other factors, it will not apply the voice interaction design strategy for the elderly to intelligent wearable devices in a fixed mode. However, it still provides optimization direction for the applications of intelligent voice devices in the aging industry and some suggestions for developing intelligent wearable devices equipped with a voice interaction system. In the future, it would have more design explorations from theoretical perspectives of acoustics, linguistics, multimodal interaction, etc. We also plan to choose a specific Chinese

elderly exercise to carry out targeted voice interaction system design and application practice research to further verify the current voice design model and make more relevant research.

REFERENCES

Alam, F. and Riccardi, G. (2014). Fusion of acoustic, linguistic and psycholinguistic features for speaker personality traits recognition. *In 2014 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 955–959.

Esfahani, B.K. and Ganji, D. (2021). The importance of product language: an exploratory study of smartwatches for remote healthcare. *In 12th International Conference on Applied Human Factors and Ergonomics*, pp. 67–73.

Harris, R. A. (2004). Voice interaction design: crafting the new conversational speech systems. *Elsevier*.

Ji, W., Liu, R. and Lee, S. (2019). Do drivers prefer female voice for guidance? An interaction design about information type and speaker gender for autonomous driving car. *In International Conference on Human-Computer Interaction*, pp. 208–224.

Jiang, H.Y. (2020). Research on interaction design method from the perspective of context theory (Master thesis). University of Science and Technology Liaoning.

Niu, L. (2021). Assistant or friend? Research on users' preferences of intelligent voice products — From the perspective of gender stereotypes(Master thesis). Zhejiang Gongshang University.

Peng, J., Li, J., Li, W., Chen, T.T. and Liu, H. (2017). Development status and Prospect of intelligent wearable devices. *West Leather*, 39(16), pp.116.

Pradhan, A., Mehta, K. and Findlater, L. (2018). "Accessibility came by accident" use of voice-controlled intelligent personal assistants by people with disabilities. *In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–13.

Qin, X. (2021). Research on the development status and countermeasures of intelligent elderly care under the background of aging. *Business & Economy*, 10, pp. 39–41.

Sarkar, S. and Chakrabarti, D. (2021). The perception and acceptance of wearable fitness devices among people and designing interventions for prolonged use. *In 12th International Conference on Applied Human Factors and Ergonomics*, pp. 94–101.

Silnitskaya, A. S. and Gusev, A. N. (2013). Character and temperamental determinants of prosodic parameters of natural speech. *Psychology in Russia*, 6(3), pp.95.

Sun, H.Y. and Wu, Y.X. (2017). The scenario interaction design in mobile applications. *Design*, 10, pp. 104–105

Turk, M. (2014). Multimodal interaction: A review. *Pattern recognition letters*, 36, pp. 189–195.

Wang, P.K. (2019). Study of voice interaction based on elderly companion robot (Master thesis). Beijing University of Posts and Telecommunications.

Yang, F., Pan, Y. and Zou, Z.Y. (2019). Patterns and determinants of physical activity of elderly people in China. *China Sport Science and Technology*, 55(10), pp. 10–21.

Zhou, J. and Zhou, M. (2021). Sentiment analysis of elderly wearable device users based on text mining. *In 12th International Conference on Applied Human Factors and Ergonomics*, pp. 360–365.