

Virtual Assistant Interaction Design for the Elderly Based on Emotional Theory

Shuangqian Liu¹ Yi Li² Rui Zhang³

^{1,2,3} Hubei Institute of Fine Arts, Wuhan, Hubei, China

ABSTRACT

Objective: This study applies emotional design theory to the design of elderly virtual assistant voice interaction, explores the influence of elderly virtual voice assistant timbre characteristics on the elderly's liking degree, and uses it as a basis for the design of elderly virtual assistant and interaction system. **Method:** Through a combination of quantitative and qualitative methods such as literature research method, Wizard of Oz survey method and Likert scale, it obtains the influence of voice assistant anthropomorphic features on the elderly's liking degree, and conducts elderly needs analysis and design point transformation for the design of elderly virtual assistant and interactive system. **Conclusion:** Elderly users are more fond of the female virtual assistant image who is enthusiastic and active, lively and lovely, with high interaction degree, featured with a bright and lively voice trait. Through the research and analysis, the functions of the elderly virtual assistant include four sections: medical management, diet and sports management, entertainment and companionship, and intelligent community.

Keywords: Interaction design, Population aging, Emotional theory, Virtual assistant.

1. INTRODUCTION

Population aging has become a serious problem faced by the entire world. The elderly are prone to loneliness, pessimism, negativity and other psychological emotions and states due to the decline of organism functions and major changes in social roles. With the continuous development of technology, human-computer interaction has also gained remarkable achievements.

Voice interaction has low learning cost and is suitable for the elderly with declining perception and learning ability. Voice interaction contains rich human emotion while carrying information, giving the majority of the elderly emotional culture transmission. Through conversational AI technology and image processing technology, it presents a new communication state in the human-computer relationship and establishes trust and intimacy with users. The virtual assistant can be given anthropomorphic features by setting its age, gender, appearance, speed of speech and voice tone, etc. This study takes virtual assistants as the research object to analyze the influence of voice features on the elderly's liking degree.

2. CURRENT STATUS OF CHINESE AND INTERNATIONAL RESEARCH

In terms of voice interaction, Chinese voice technology is in a continuous optimization stage, with recognition accuracy rates reaching over 97% [1], and emotional understanding and expression has been introduced into the field of cognitive science to enhance the anthropomorphic communication capabilities of artificial intelligence.

At present, in terms of research on theoretical issues related to the application of voice interface technology for elderly intelligent products, Jia Guozhong believes that the current direction of the relevant research field of elderly intelligent products mainly focuses on the design of elderly intelligent companion products [2] and the design of wakeup words and voice functions; in terms of research on the design of voice characters, the elderly generally tend to prefer a more gentle and quiet female image. Wang Pankai analyzes the various cognitive loads of the elderly in voice interaction activities, and proposes design methods

such as using self-explanatory language for interaction, automatically adjusting the voice, and adding the special language of the elderly for design interaction in the interaction design process of the elderly language [3]. Dou Jinhua also points out the general trend that intelligent voice interaction design based on voice user emotional experience and user cognitive interaction mode design is another important development direction of intelligent human-computer voice interaction design in the future. The existing relevant research results lack interface design strategies for smart home voice users for various specific scenario contextual needs. [4]

In terms of the degree of anthropomorphism of voice assistants, Liao Qinglin believes that there is a positive correlation between the degree of product perception anthropomorphism and emotional interaction, and under the same conditions, female images are more anthropomorphic compared to male images, and user satisfaction increases with the increase of anthropomorphism. However, care should be taken to avoid "The Uncanny Valley Effect". [5]

In a study on the acceptance of voice assistants by the elderly, Linda Wulf et al. study the acceptance of voice-only interactions by the elderly and shows that voice-based system interactions are recognized and widely accepted by the elderly, providing a simple way for the elderly to interact with modern technology [6].

In the study on the effect of the degree of anthropomorphism of voice assistants on user satisfaction, Romportl's experiments show that most people tend to prefer a voice with a higher degree of naturalness [7]. Current foreign

theoretical studies on the voice emotional interaction aspects of human-computer are more concerned with the basic attributes of voice such as frequency, duration, tone quality and intelligibility [8], mostly focusing on the two aspects of voice recognition technology and voice emotion computing. Niculescu et al. study and analyze the impact of social robots' intonation, sense of humor, and empathy on the user's social experience [9]. Lee et al. conclude that emotional experience and deep companionship are very critical influencing factors on whether smart home products can improve user satisfaction [10]. However, few studies have been conducted on intelligent assistants and user emotional interaction models.

In summary, current research on voice interaction design has focused on the cognitive difficulties of the elderly. For the anthropomorphic research of voice assistants, both China and foreign countries have focused only on the younger groups, and the research on the voice design of virtual assistants for the elderly is still in a vacant state.

3. THEORIES AND RESEARCH METHODS FOR THE DESIGN OF VOICE FEATURES OF ELDERLY VIRTUAL ASSISTANTS

3.1 Emotional Design Theory

The hierarchical theory of user emotion has been explained in detail in the book Emotional Design. According to author Norman, this theory includes three levels of instinct, behavior, and reflection [11], as shown in "Figure 1" and "Figure 2".

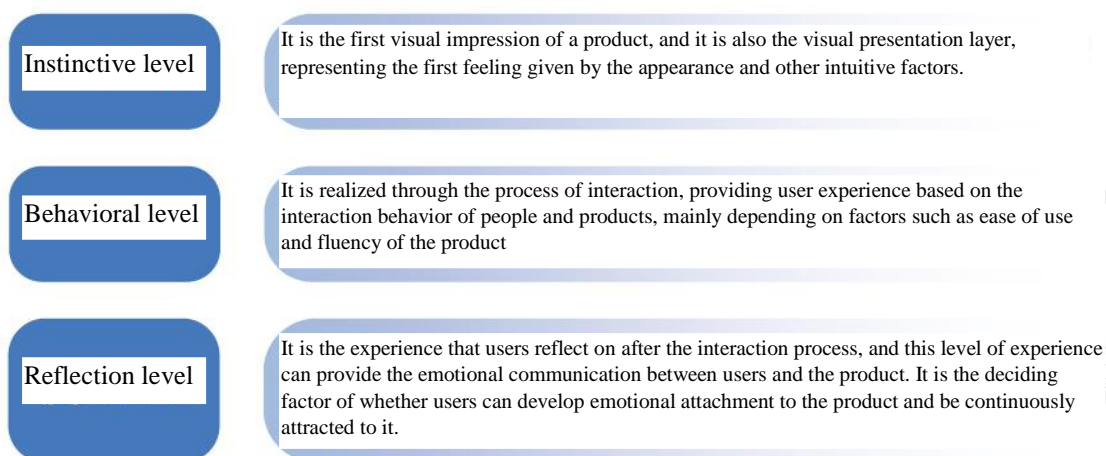


Figure 1 Emotional theory illustration 1. (Source: drawn by the authors)

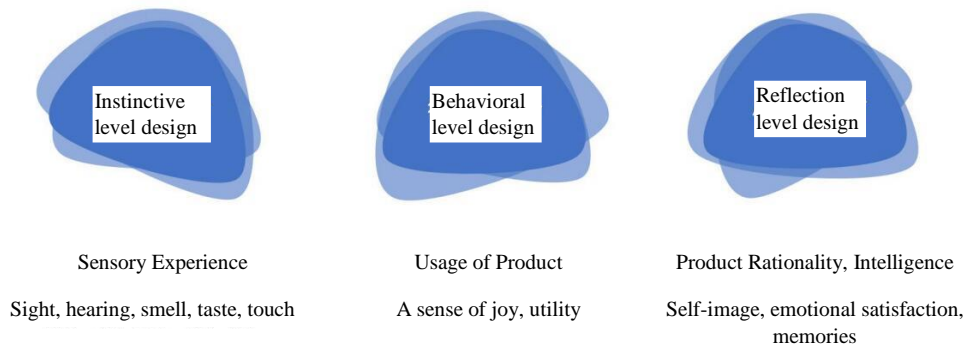


Figure 2 Emotional theory illustration 2. (Source: from the Internet)

The elderly compensate for missing emotions by re-establishing and accomplishing new goals in life, which is the principle of emotional compensation. It achieves emotional compensation by satisfying the implicit needs of the user to achieve a reflective level of experience. When facing the elderly user group, efforts should be made to get rid of the rigid one-way interaction mode and give the elderly emotional care in a more lively and humanized interaction mode, so as to establish a new type of emotional attachment to achieve emotional compensation.

Language is the most direct way for people to communicate and it is also one of the main ways to make interaction emotional. The development of voice interaction has changed from one-way output communication to natural two-way communication, and is widely used. It is possible to reduce the learning cost of interaction through technical factors, get rid of the rigid one-way interaction mode, give emotional care to the elderly with a more humanized interaction mode, and establish emotional attachment to achieve compensation.

3.2 Impact of Virtual Assistant Timbre on Emotional Design

Emma Lodeiro, a professor of communication, suggested that "When people hear a voice, they have different impressions of the voice according to their life experiences and unconsciously construct a specific image for the voice" [12]. The anthropomorphic features and emotional information contained in the voice are much larger than the content of the voice information itself. Anthropomorphism is not only about providing a friendly tone of voice, but its three-dimensional, flexible and dynamic character settings can also deepen users' understanding and emotional

engagement, increasing users' stickiness and dependence. Providing personality-filled characters, good roles and character settings can help the emotional connection between users and the product, and is also an important way to build users' emotional connection at the reflective level, thus achieving emotional compensation, so that the negative emotions generated in the life of the elderly can be compensated.

In *Connecting Voices*, Clifford Nass, a professor of communication at Stanford University, has theorized in detail about the social and physical properties of voice and the various stereotypes inherent in society at large about voice, respectively. He argues that in the human subconscious, women's voices are amiable and give a sense of reliability, while men's voices tend to be more authoritative [13]. For the study of voice trustworthiness, scholars from Soochow University found that for people's first exposure to timbre, female timbre is perceived to be much more credible than male's, and in the experiment, most testers show a preference for female timbre [14].

According to the theory of emotional design, the product should meet the emotional needs while paying attention to personal awareness, and shape the character through the difference of timbre to respond to the preferences of different user groups. Providing characters with different settings and characteristics for different user groups can better connect with users emotionally.

3.3 Wizard of Oz Survey Method

The Wizard of Oz test method, also known as the "WOZ" test method, is often used in the prototyping of voice interactions to quickly test the system at different stages of development and save development costs. The experimenter usually

narrates the background of the experiment and the task requirements as a "wizard" and operates the product system in an adjacent room without the knowledge of the subject. This method has the advantages of low cost, easy set-up, and timely feedback.

In studying the acceptance of products with voice interaction by the elderly, Francois Portet et al. found that voice interaction has great potential to improve the daily lives of the elderly through the Wizard of Oz test method [15].

4. EXPERIMENTAL STUDY OF VIRTUAL ASSISTANT VOICE

It conducts experiments based on virtual assistant timbre and anthropomorphic features to explore the correlation between different virtual assistant voice features on the liking degree of elderly users and the correlation between different

Table 1. Table of anthropomorphic features of the voice samples (Source: self-made)

Experimental samples	Gender	Voice traits	Personality keywords
No.1, No.2	Female/Male	Gentle and wise	Mature and prudent, kind and affable
No.3, No.4	Female / Male	Clean and fresh	Gentle and intellectual, positive
No.5, No.6	Female/Male	Bright and lively	Dynamic and lovely, witty and humorous

4.1.2 Simulation Experiment

To ensure the validity of the experimental samples, three experts who are more knowledgeable about voice assistants are invited to conduct the simulation experiment to test the

Table 2. Sample dialogue script I (Source: self-made)

Sample	Script
User	What's the weather today?
No.1, No.2	Today it is cloudy to drizzly in Wuhan, with a minimum temperature of 10°C and a maximum of 16°C. Do remember to bring an umbrella and please don't catch a cold when the temperature drops.
No.3, No.4	Today it is cloudy to drizzly in Wuhan, with a minimum temperature of 10°C and a maximum of 16°C. Please wear more clothes to avoid getting chilled. By the way, do not forget to go out with an umbrella.
No.5, No.6	Today in Wuhan it will be cloudy to drizzly, with the possibility of rain, so please remember to take an umbrella when you go out. The lowest temperature is 10°C, the highest temperature is 16°C, and it is much colder than yesterday, so remember to put on more clothes.

Table 3. Sample dialogue script II (Source: self-made)

Sample	Script
User	Can you play a song for me?
No.1, No.2	Okay, I guess you want to hear this song "The East Is Red", see if I'm right?
No.3, No.4	Okay, let me recommend you a song "The East Is Red", see if I know you well?
No.5, No.6	No problem. How about I sing you a song called "The East Is Red"? I'm a small library of songs, and I can sing anything.

tone timbre and user liking feedback, and summarizes the design strategy of voice interaction anthropomorphism.

4.1 Experimental Design

4.1.1 Experimental Group Setup

It took the mainstream intelligent voice assistants on the market for language role feature analysis and keyword extraction; to avoid the influence of gender differences on users' subjective perception evaluation, two different tones of male and female were set for each type, keeping the script, tone, and volume unchanged, for a total of six groups of voice experimental samples for voice interaction ("Table 1"). It used Likert scale to collect perceptual cognitive feedback from users, observed and recorded subjects' thoughts during the test, and combined with the data for analysis.

authenticity and fluency of the voice dialogue samples one by one, and to judge whether their dialogue contents show the characteristics of the corresponding voice samples. ("Table 2", "Table 3", "Table 4")

Table 4. Sample Dialogue Script III (Source: self-made)

Sample	Script
User	Can you chat with me?
No.1, No.2	Of course. I'm happy talking to you too.
No.3, No.4	Sure thing. I guess you've come across something happy, can you share it with me?
No.5, No.6	I love to talk with you. Do you have anything happy to tell me? I'm looking forward to it.

4.1.3 Wizard of Oz Voice Feature Experiment

It used a random sampling method to select eight elderly users who had experience using smart products as the research subject population. To ensure the breadth of the research, the subject population covered different occupations, genders and age spans.

At the beginning of the experiment, the experimental operator briefly introduced the process to the subjects so that they would think that the voice interaction was real and in a calm and

relaxed state. The experiment was later conducted with the experimenter's prompting. Another experimenter played the role of "wizard" to control the voice assistant in order to complete the experiment process.

The subjects rated the sample on a five-point Likert scale; the scores were then collected and tallied. After all the experiments of the six groups were completed, the subjects were interviewed individually and asked to rate and rank the overall performance of each group and to give an open-ended account of their subjective perceptions of the test sample and the reasons for them. ("Table 5")

Table 5. Excerpts from the design of the Likert scale (Source: self-made)

Dimension	Statement
Satisfaction	I think this voice assistant is taking care of my emotions
	I think this voice assistant is gentle
	I think this voice assistant can be trusted
	I think this voice assistant is attentive
	I think this voice assistant is reliable
	I will recommend this voice assistant to my friends
	I feel this voice assistant will communicate with me in the way I expect
	I think this voice assistant is fun to use
	I think I need to have this voice assistant

4.2 Discussion of Experimental Results

It integrated the Likert scale for data analysis, and the mean value of each index of each voice

assistant was selected for comparative analysis. The results are as follows ("Table 6").

Table 6. Analysis of the results of the Likert scale (Source: self-made)

Grouping	No.1	No.2	No.3	No.4	No.5	No.6
I think this voice assistant is taking care of my emotions	3.75	3.625	3.5	3.125	4.25	4.375
I think this voice assistant is gentle	3.625	3.125	2.875	3.5	4.5	4.125
I think this voice assistant is kind	4.25	3.375	3.25	3.25	4.125	4.375
I think this voice assistant can give me the feeling of caring	3.625	3.25	3.5	3.5	3.75	4
I think this voice assistant is dynamic	3.375	3.25	4.25	4	3.625	3.875
I think this voice assistant can be trusted	4.25	3.875	3.78	3.375	4	3.75
I think this voice assistant is attentive	3.75	3.375	3.5	3.625	3.875	3.875
I think this voice assistant is reliable	4.125	4.125	3.25	3.625	3.75	4
I am satisfied with this voice assistant	4.125	3.625	3.625	3.5	4.375	4.375

Grouping	No.1	No.2	No.3	No.4	No.5	No.6
I will recommend this voice assistant to my friends	3.75	3.625	3.25	3.375	4	3.625
I think this voice assistant is fun to use	3.25	3.25	3.25	3.5	4.375	3.5
I feel this voice assistant will communicate with me in the way I expect	3.125	3.125	3.375	3.5	3.75	3.75
I think I need to have this voice assistant	3.75	3.625	3.5	3.125	3.75	4
This voice assistant is a pleasure to use	3.875	3.5	3.625	3.5	4.5	4.25
Mean	3.76	3.48	3.47	3.46	4.04	3.99

The experimental results show that the scores of female samples are generally higher than those of male samples; and among the female samples, those with bright and lively, cute and witty humor have the highest scores. It can be seen that elderly users are more fond of the female virtual assistant image who is enthusiastic and active, lively and lovely, with high interaction degree, featured with a bright and lively voice trait, and female images are going to be more popular.

This conclusion prepares the theory and data for the later design of the visual image of virtual assistants.

5. DESIGN PRACTICE OF VIRTUAL ASSISTANTS FOR THE ELDERLY

Based on the findings of the previous empirical and qualitative studies, the authors design the elderly virtual assistant features to use AR virtual reality technology for more intuitive emotional communication. The design provides four functional sections of medical management, healthy management, entertainment and companionship, and community service, and the interaction model is based on voice interaction, supplemented by interface interaction to achieve multi-channel interaction model.

5.1 *Virtual Assistant Design*

5.1.1 *Image Design*

Combined with the emotional theory, the virtual assistant image is set to simulate a real person by setting the characteristics of tone, timbre, and movement to interact and transmit information, giving humanistic care to the elderly and establishing emotional connection.

The results of the previous quantitative experiments show that the more popular virtual assistant voice image characteristics for the elderly are lively, bright and lovely female characters.

According to this role characteristic, the identity of the virtual assistant is positioned as a granddaughter in this design. In order to avoid "The Uncanny Valley Effect" caused by the visual image being too close to the real person, an animated anthropomorphic image is chosen for the virtual assistant visual image design. The diversity of user groups means that the service should not be uniform. It is the trend of the times to segment users and study their individual needs [16]. In order to cater to the preferences of different elderly people, four different image modelings are designed around the image features. The final visual image is shown in "Figure 3" and "Figure 4".



Figure 3 Virtual assistant visual image drawing. (Source: drawn by the authors)

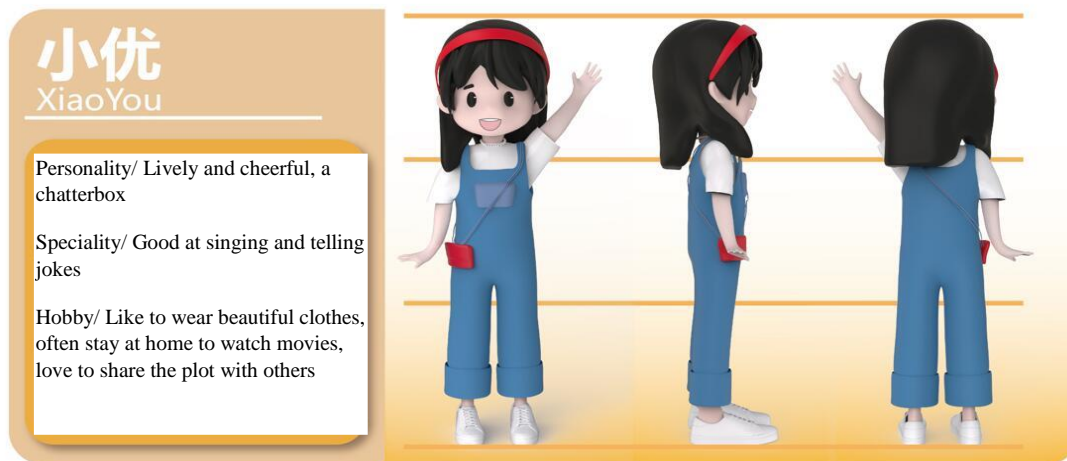


Figure 4 Virtual assistant character setting diagram. (Source: drawn by the authors)

5.1.2 Voice Trait Design

The design example dialogues based on basic character traits and character positioning are shown in the table ("Table 7"), and the character image traits are reflected by setting personalized phrases.

Table 7. Dialogue table of XiaoYou's voice image example (Source: self-made)

Character	Daily dialogues	Help-seeking dialogues
XiaoYou	Let me spend this wonderful day with you.	Don't worry, the fee is already paid for you.
	It's a beautiful day, why don't you go out and feel the sunshine?	I'll take care of it, don't worry.
	Do you want to see my Tai Chi? I'm good at it.	I'm happy to help you. You can call me if you have any questions.

"Let me spend this wonderful day with you" reflects the positive, optimistic and lively character. "Don't worry", "Leave it to me" and "I'm glad I can help you" reflect the character's reliable, professional, yet warm personality.

The tone is based on previous empirical studies and character settings, and the tone of a young girl

around 8-12 years old is chosen; the overall tone is bright and clear, and lively and light, which is in

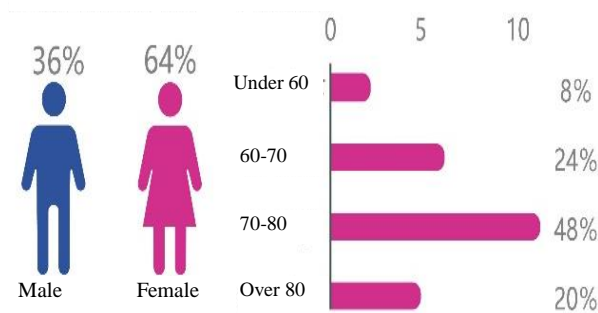
line with the character settings and brings a pleasant emotional experience.

5.2 Service Prototype Design

A research was conducted to design a complete virtual interaction system for the elderly. Through

questionnaires and user interviews, 25 elderly people were visited to clarify their needs and behavioral habits of the virtual assistant, and 121 valid messages were summarized and analyzed to find out the keywords, as shown in "Figure 5".

Gender proportion of interviewees / Age information



Summary of interview information

User needs	Keyword extraction
Want someone to talk to me	Conversation/communication
Sometimes I always forget things and want to be reminded	Reminder of events
Fail to do what I want to do easily	Difficult to operate
Want to know more about my health condition	Health management
I hope it can help with my health	Convenient life
Poor eyesight, can't see the screen, want to listen more with ears	Voice communication
Hope it can be interesting	Recreational activities
Want to be able to relax and enjoy myself	

Figure 5 Elderly virtual assistant demand research and analysis diagram. (Source: made by the authors)

It combines the user demand points to simulate and envision the daily life behaviors and unexpected situations of elderly users, explore the users' emotions and thoughts, so as to find the design points transform the innovative design points, and create the service prototype diagram, as shown in "Figure 6".

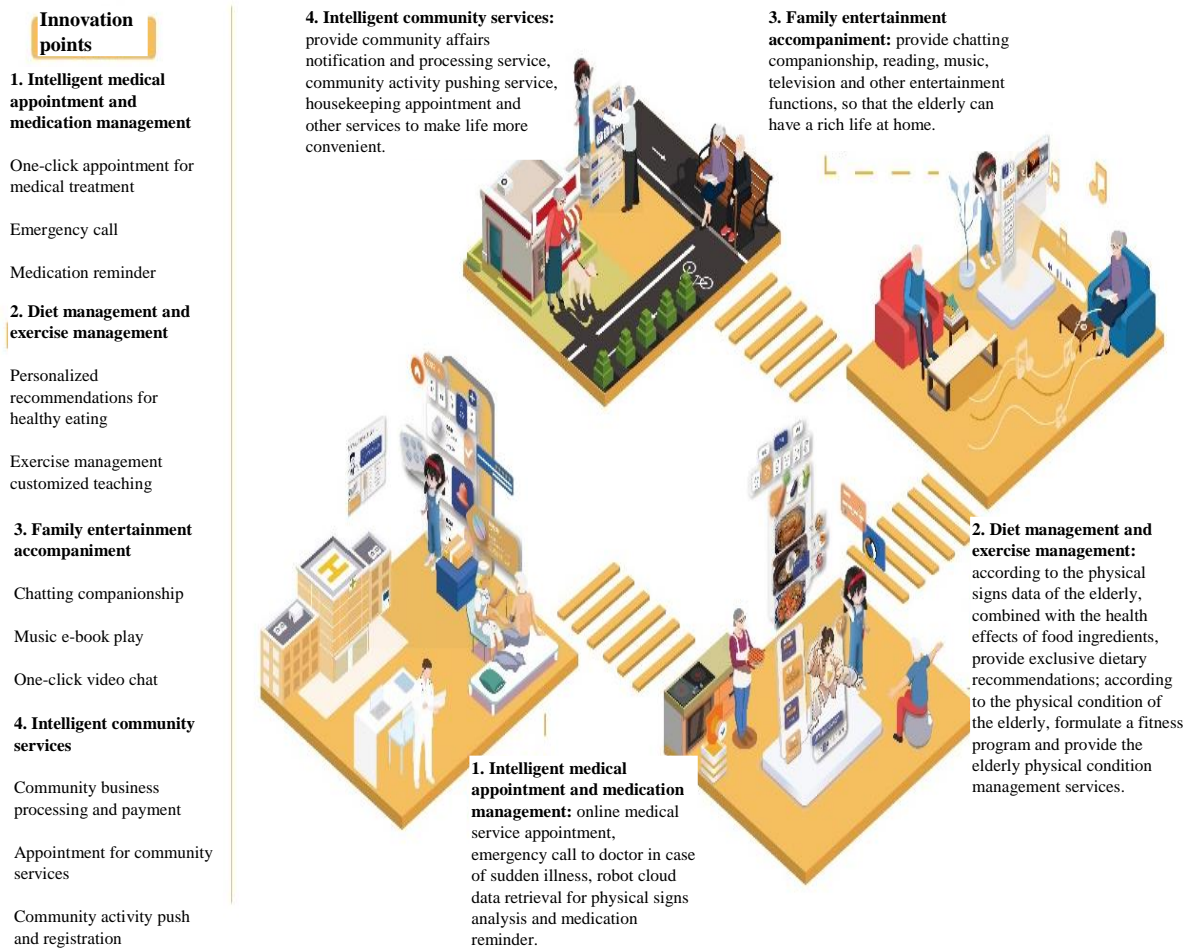


Figure 6 Elderly virtual assistant service prototype diagram. (Source: made by the authors)

5.3 Framework Prototype and Usage Scenario Construction

This design is based on the cognitive characteristics of the elderly to layout the content and creates a framework prototype ("Figure 7"). It takes into account the cognitive characteristics, concentration and usage habits of the elderly, and designs the health care and entertainment community sections. Compared with the traditional interaction method, this is an interaction method with voice interaction as the focus, and the interface interaction part is used as a secondary interaction to present the information, adopting the square tabular interface design preferred by the elderly [17]. The area of the operation area is larger than other interfaces on the market, and the overall content elements are clearer and larger, and the operation

steps are reduced. The enlarged and simple operation interface can improve the visibility and ease of operation for elderly users and make the operation process simple and clear. The function area animation and text presentation are in parallel, being more intuitive and convenient, reducing the cognitive burden of the elderly.

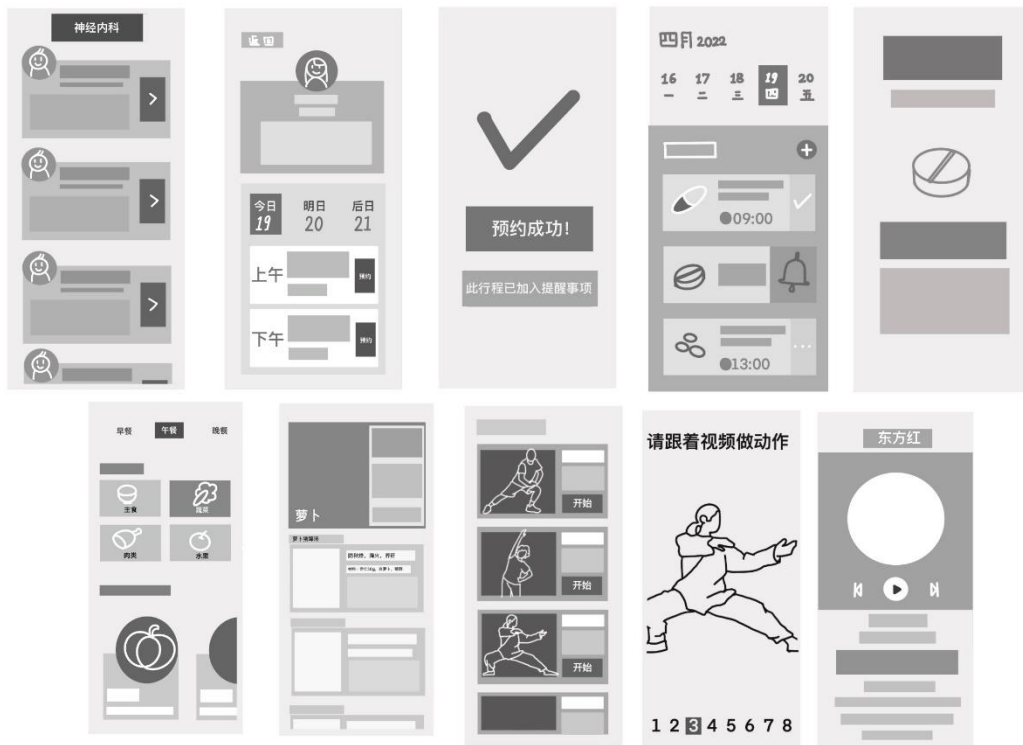


Figure 7 Framework prototype diagram. (Source: made by the authors)

The virtual assistant visual image and usage interface are presented through virtual reality technology, with the home scene as the main usage scenario. Compared with the traditional virtual assistant that interacts with the medium of screen, it is more interactive and conducive to the active interaction of the elderly, making it possible to obtain the satisfaction of implicit needs to reach the reflective level of experience, generate emotional connection and provide caring services with a lower desire for exploration. The usage is shown in "Figure 8" and "Figure 9".



Figure 8 Scene construction effect diagram 1. (Source: made by the authors)



Figure 9 Scene construction effect diagram 2. (Source: made by the authors)

6. CONCLUSION

Based on the theory of emotional design, this paper uses the principle of emotional compensation to establish the emotional attachment of the elderly through virtual assistant interaction, and researches and designs the voice characteristics of virtual assistants for the elderly to realize the emotional experience at the reflective level. It uses the Wizard of Oz research method and Likert scale for sample experimental analysis to extract keywords for the anthropomorphic features of the elderly's favorite voice assistant. The experiment provides a design basis for elderly voice interaction, making the voice

design more in line with the preferences of the elderly and enhancing the interaction experience. In the elderly AR virtual assistant design example, it further designs the interaction system and interface through the analysis of elderly behavior habits, which provides some reference value for the design of elderly virtual assistants.

AUTHORS' CONTRIBUTIONS

Shuangqian Liu analyzed and compiled the experimental data of the paper, Yi Li sorted out the overall logic of the paper, and Rui Zhang revised and edited the paper.

REFERENCES

- [1] How "Internet+Nursing Services" Make Precise Efforts-Ministry of Civil Affairs of the People's Republic of China (mca.gov.cn) [EB/OL].
<http://www.mca.gov.cn/article/xw/mtbd/201808/20180800010649.shtml>,2018-08-20. (in Chinese)
- [2] Jia Guozhong. Research on voice interaction design of smart speakers for the elderly [D]. South China University of Technology, 2018. (in Chinese)
- [3] Wang Pankai. Research on the design of voice interaction for elderly companion robots [D]. Beijing University of Posts and Telecommunications, 2019. (in Chinese)
- [4] Dou Jinhua, Qi Ruoxuan. Elderly-Adaptability Voice User Interface Design Strategy of Smart Home Products Based on Context Analysis [J]. Packaging Engineering, 2021, 42(16): 202-210. (in Chinese)
- [5] Liao Qinglin, Wang Mei, Feng Zhan. Voice Interaction Design of Smart Home Products Based on Emotional Interaction [J]. Packaging Engineering, 2019, 40(16): 37-42+66. (in Chinese)
- [6] L. Wulf, M. Garschall, J. Himmelsbach, and M. Tscheligi. Hands Free -Care Free: Elderly People Taking Advantage of Speech-only Interaction[J].In Proceedings of the 8th Nordic Conference on Human-Computer Interaction:Fun,Fast,Foundational.2 014,14:203–206.[J]
- [7] ROMPORTL J. Speech Synthesis and Uncanny Valley [C]//ROMPORTL J. Text, Speech and Dialogue.Brno: Springer International Publishing, 2014: 595—602.[J].
- [8] Mao Xia, Xue Yuli. Human-computer emotional interaction [M]. Beijing: Science Press, 2011. [J] (in Chinese)
- [9] Niculescu A, Van Dijk B, Nijholt A, et al. Making Social Robots More Attractive: The Effects of Voice Pitch, Humor and Empathy[J]. International Journal of Social Robotics, 2013, 5(2): 171-191.
- [10] Lee B, Kwon O, Lee I, et al. Companionship with smart home devices: The impact of social connectedness and interaction types on perceived social support and companionship in smart homes[J]. Computers in Human Behavior, 2017, 75: 922-934.
- [11] DONALDAN. Design Psychology 1 — Everyday Design [M]. Beijing: Citic Press, 2015. DONALDAN. Design Psychology 1: Everyday Design [M]. Beijing: CITIC Press, 2015. (in Chinese)
- [12] Emma Rodero. The Spark Orientation Effect for Improving Attention and Recall[J]. Communication Research,2019,46(7)
- [13] Chen Xiaoliang. Why are voice assistants mostly female voices [N]. Beijing Science and Technology News,2015-11-23(053) (in Chinese)
- [14] Ding Yitong. The influence of voice trustworthiness stereotypes on trust [D]. Soochow University,2018 (in Chinese)
- [15] Portet F, Vacher M, Golanski C, et al. Design and Evaluation of a Smart Home Voice Interface for the Elderly: Acceptability and Objection Aspects[J]. Personal and Ubiquitous Computing, 2013, 17(1): 127-144.[J]
- [16] Wang Xiaobing, Jiang Min, Zhang Rui. Research on the construction of a wellness service platform based on personalized demand [J]. Art Science and Technology, 2019,32(07):68. (in Chinese)
- [17] YAO Jiang, FENG Bing. Interface Interactive Design of Information Products for the Elderly from the Perspective of Experience[J]. Packaging Engineering, 2015, 36(2): 67-71.