

Exploration of the Path to Intelligent Computer Operation for Visually Impaired People

Xie Yingting, Wang Kunyu, Li Guanchi, Yang Zeyuan

School of Management Science and Engineering, Anhui university of finance and economics Email address: 20213908@aufe.edu.cn

Abstract—With the rapid development of information technology, computers have become an important tool for people to obtain information and communicate. However, visually impaired people face many obstacles when using computers. This article aims to explore the path of intelligent computer operation for visually impaired people. By analyzing existing technical solutions, system designs and application cases, it provides visually impaired people with more convenient and efficient computer operation methods, and improves their quality of life and social participation.

Keywords—Visually impaired people, Accessibility technology, OpenCV image recognition technology, Optical character recognition technology.

I. INTRODUCTION

The rapid development and popularization of computer technology have greatly promoted the process of social informatization, bringing unprecedented convenience to people's life, study, work and entertainment [1]. However, for the special group of visually impaired people, the wide application of computer technology has not simultaneously enhanced their ability to acquire and operate information. The traditional computer operation mode relies highly on visual feedback. Users need to complete the interaction with the computer through visual elements such as text, ICONS and buttons on the screen. However, visually impaired people, due to their visual impairments, cannot see the content on these screens as intuitively as ordinary people, which leads to huge challenges for them when operating computers.

Visually impaired people often get into trouble when using computers because they cannot accurately recognize the information on the screen. They may not be able to independently open applications, browse web pages, edit documents or perform other basic computer operations. This not only seriously affects their learning efficiency and working ability, but also limits their participation and autonomy in the information society. Therefore, exploring the path of intelligent computer operation for visually impaired people is of great practical significance for breaking this technical barrier and improving the information literacy and quality of life of visually impaired people[2].

The research on the path of intelligent computer operation for visually impaired people aims to provide visually impaired people with more convenient and efficient computer operation methods through technological innovation and system optimization. This can not only help them overcome the operational difficulties brought about by visual impairments, but also promote the development of barrier-free technologies and drive the all-round progress of social informatization. Meanwhile, this research also helps to enhance society's attention and support for visually impaired people, promote social inclusiveness and diversity, and contribute to building a more harmonious and equal information society.

II. COMPUTER OPERATION REQUIREMENTS FOR VISUALLY IMPAIRED PEOPLE

Visually impaired people often get into trouble when using computers because they cannot accurately recognize the information on the screen. They may not be able to independently open applications, browse web pages, edit documents or perform other basic computer operations. This not only seriously affects their learning efficiency and working ability, but also limits their participation and autonomy in the information society[3].

Therefore, in-depth research on the computer operation needs of visually impaired people and exploration of technical solutions to meet these needs are of great practical significance for breaking this technical barrier and improving the information literacy and quality of life of visually impaired people. In order to better meet the computer operation needs of visually impaired people and enhance their convenience and efficiency in using computers, the following is an in-depth analysis of their needs from multiple aspects:

A. Requirements for Obtaining screen content

Visually impaired people cannot directly obtain the information on the screen through vision, so they urgently need an effective way to understand the screen content. This includes various interface elements such as text, images, buttons, ICONS, etc. For example, when reading documents, they need to know the textual content in the documents; When browsing web pages, it is necessary to understand elements such as links, pictures and advertisements on the web pages. To meet this demand, technologies such as screen readers have emerged. They can convert the information on the screen into voice or haptic feedback to help visually impaired people understand the screen content[4]. However, the functional requirements of screen readers are far more than this. They also need to support multiple languages to meet the language needs of people with different visual impairments. At the same time, functions such as adjusting the speaking speed and intonation are also indispensable to ensure that visually



impaired people can obtain information according to their own auditory habits.

B. Interface Element Operation Requirements

When visually impaired people use computers, they need to operate various elements on the interface, such as clicking buttons, inputting text, and selecting menus. However, because they cannot see the positions and states of interface elements, they often have difficulty operating accurately. Therefore, they need a technology that can precisely locate interface elements and simulate operations. Technologies such as pywinauto GUI automation tools can achieve precise positioning and operation of interface elements, thereby meeting the operation requirements of interface elements for visually impaired people. However, the challenges of operating interface elements are not limited to this. Issues such as dynamic interfaces and complex layouts may also cause operational difficulties for visually impaired people.

Therefore, future technological development needs to pay more attention to solving these problems in order to enhance the convenience and efficiency of visually impaired people in using computers.

C. Voice Interaction Requirements

Voice interaction is an important way for visually impaired people to use computers. They can control computer operations through voice commands, such as opening applications, browsing web pages, editing documents, etc. Voice interaction technology needs to have high accuracy and robustness to adapt to voice input in different accents, speaking speeds and noisy environments[5].

Meanwhile, in order to meet the natural interaction needs of visually impaired people, voice interaction technology still needs to continuously improve its naturalness and user experience. In recent years, with the continuous development of deep learning technology, the accuracy of speech recognition technology has been significantly improved, which provides better support for visually impaired people to use speech interaction technology.

D. Personalized Setting Requirements

Different visually impaired people may have different usage habits and needs, so they hope that the computer system can provide personalized Settings. For instance, adjust the voice speed, font size, screen contrast, etc., to suit your visual and auditory abilities. Technologies such as personalized Settings and learning modules can meet this demand[6].

They allow users to customize system Settings according to their preferences and needs, and automatically adjust parameters through machine learning technology. However, the demand for personalized Settings is not limited to this. Future technological development also needs to pay more attention to meeting the diverse needs of visually impaired people, such as providing more customization options and supporting more assistive functions.

E. Requirements for System Stability and Reliability

Visually impaired people often rely more on the stability and reliability of the system when using computers. Once the system malfunctions or errors occur, they may not be able to detect the problem in time and fix it, resulting in operational interruption or data loss. Therefore, they need a stable and reliable computer system that can operate stably for a long time and promptly alert users or automatically repair faults when they occur.

To meet this demand, technologies such as system stability and reliability guarantee mechanisms have emerged. They enhance the stability and reliability of the system through technical means such as redundant design, fault prediction and health management.

However, with the continuous development of computer technology, the complexity of the system is also constantly increasing, which puts forward higher requirements for the stability and reliability of the system. Therefore, future technological development needs to pay more attention to improving the stability and reliability of systems to ensure that visually impaired people can use computers with greater peace of mind.

III. TECHNICAL SOLUTIONS

A. Research on Image Recognition and Interface Element Location Technology

OpenCV, as a powerful computer vision library, can process and recognize images efficiently. In the research on intelligent computer operation for visually impaired people, by training the OpenCV model, it can recognize elements such as text, ICONS, and buttons on the screen, and convert the recognition results into voice output or haptic feedback to help visually impaired people understand and operate the computer interface[7].

Image recognition is carried out by using computer vision libraries such as OpenCV. Methods such as feature extraction, template matching or deep learning are adopted to locate and recognize interface elements such as buttons, text boxes and ICONS in computer screenshots. At the same time, attention is paid to improving the accuracy and efficiency of image recognition to meet the needs of different visually impaired people.

B. Research on GUI Automation Script Writing Technology

PYWINAUTO is a Python library for Windows GUI automation, which can simulate operations such as mouse clicks and keyboard inputs to achieve automatic control of the computer interface. In the research, it is used to construct automated scripts[8]. Based on the screen element information recognized by OpenCV, corresponding operations are automatically executed to achieve precise control of the computer interface and help visually impaired people complete complex operation tasks.

Based on GUI automation tools such as pywinauto, scripts are written to simulate GUI operations such as mouse clicks and keyboard inputs. Based on the interface elements located according to the image recognition results, the corresponding operations are automatically executed to meet the computer operation requirements of visually impaired people[9]. Pay attention to aspects such as the writing norms of scripts,



optimization algorithms, and exception handling to improve the stability and reliability of automated operations.

C. Research on Voice Interaction and Command Recognition Technology

Integrated speech recognition technology enables visually impaired people to control computer operations through voice commands. The conversion from speech to text is carried out by using the existing speech recognition model, and an appropriate speech instruction set and parsing algorithm are designed to enable the system to accurately recognize speech instructions and trigger the corresponding GUI automation scripts. Meanwhile, attention should be paid to improving the accuracy and robustness of speech recognition to adapt to speech input in different accents, speech rates and noisy environments.

This technology combines multiple algorithms and techniques. For instance, traditional command word recognition methods utilize the formant characteristics of speech signals[10]. Modern speech recognition technology introduces algorithms such as dynamic programming and linear predictive analysis. Deep learning technology also provides new research directions for speech interaction and command recognition technology.

D. Optical character Recognition technology

OCR technology can convert the text displayed on the screen into editable and readable text formats. By combining the capture of screen content by cameras, OCR technology can recognize the text on the screen and convert it into voice output for visually impaired people to listen to. This is particularly important for tasks such as reading documents and browsing web pages.

IV. SUMMARY AND CHALLENGES

The rapid development and popularization of computer technology have brought unprecedented convenience to people's life, study, work and entertainment. However, for the special group of visually impaired people, the wide application of computer technology has not simultaneously enhanced their ability to acquire and operate information. The traditional way of operating computers is highly dependent on visual feedback. However, visually impaired people, due to their visual impairments, cannot see the content on the screen as intuitively as ordinary people, which leads to huge challenges for them when operating computers. In order to break through this technical barrier and enhance the information literacy and quality of life of visually impaired people, it is particularly important to study the path of intelligent computer operation for visually impaired people.

The image recognition function of OpenCV can provide precise operation instructions for pywinauto. pywinauto can simulate various mouse and keyboard operations, achieve comprehensive control of the computer interface, improve the intelligent level of the system, integrate the above technologies into a complete system, and design an interactive interface that is easy to understand and operate. Reduce the operational skills requirements for visually impaired people for their use.

The system should have a high degree of customizability and scalability, and be capable of personalized Settings based on the specific needs and preferences of visually impaired people. At the same time, pay attention to the stability and performance optimization of the system to ensure a good user experience during actual use[11]. The design process includes image recognition using OpenCV, GUI automation operation using the pywinauto library, integration and testing, as well as optimization and adjustment.

At present, certain progress has been made in the research on the intelligence of computer operation for visually impaired people. By introducing technologies such as image recognition, voice interaction, and GUI automation, a more convenient and efficient computer operation method is provided for visually impaired people[12]. These technologies can not only help visually impaired people overcome the operational difficulties brought about by visual impairments, but also promote the development of barrier-free technologies and drive the all-round progress of social informatization.

However, despite these advancements, the intelligentization of computer operation for visually impaired people still faces many challenges, such as:

A. Technical bottleneck

Accuracy and efficiency of image recognition: Although image recognition technologies such as OpenCV have made significant progress, they still face challenges in terms of accuracy and efficiency when dealing with complex interfaces, dynamic elements, and low-quality images. Visually impaired people need to obtain information on the screen quickly and accurately, which puts forward higher requirements for the real-time performance, accuracy and robustness of image recognition technology [13].

The robustness of speech recognition: Speech interaction technology offers visually impaired people a convenient operation method, but the accuracy of speech recognition still needs to be improved in terms of different accents, speech rates, noisy environments, and dialect recognition. In addition, the naturalness, smoothness and user experience of voice interaction also need to be further improved to better meet the actual needs of visually impaired people.

System stability and reliability: Visually impaired people rely more on the stability and reliability of the system when using computers. However, with the continuous development of computer technology, the complexity of the system is also constantly increasing, which puts forward higher requirements for the stability and reliability of the system. How to ensure that the system remains stable during long-term operation and can be restored in a timely manner when faults occur is an important challenge currently faced.

B. User acceptance

Technical adaptation period: Visually impaired people may have a low acceptance of new technologies and need time and education to adapt to the new operation methods. How to lower the learning curve and provide intuitive and easy-to-use user interfaces and interaction methods is the key to improving user acceptance.

Personalized needs: Different visually impaired people may have different usage habits and demands. How to meet their personalized needs and provide a customized computer operation experience is also an important challenge. This requires the system to have a high degree of flexibility and configurability to meet the specific needs of different users.

C. Social Cognition and Support

Accessibility awareness: Although accessibility technology has made certain progress, society's understanding of the accessibility needs of visually impaired people is still insufficient. Enhancing public awareness and support for barrier-free technologies is an important factor in promoting the development of this field.

This requires enhancing society's understanding and recognition of barrier-free technologies through various means such as media promotion, public welfare activities, and education and training.

Policies and Standards: At present, the policies and standards concerning barrier-free technology are still not perfect, which to some extent restricts the development and application of barrier-free technology. Formulating and improving relevant policies and technical standards, and clarifying the requirements and norms of barrier-free technologies are important guarantees for promoting the development of barrier-free technologies.

Facing these challenges, we need to constantly innovate technologies, expand application scenarios, and strengthen policy and social support to jointly promote the development of intelligent computer operation for visually impaired people[14]. In the future, with the continuous advancement of technology and the expansion of application scenarios, we have every reason to believe that visually impaired people will be able to use computers more conveniently and efficiently, better integrate into the information society, and realize their self-worth.

ACKNOWLEDGMENT

Fund Project: 2024 Undergraduate Scientific Research and Innovation Fund Project of AnhuiUniversity of Finance and Economics: Based on OpenCV image recognition and pywinautoGUI research on Intelligent Computer Operation for Automatically Visually Impaired Persons (XSKY24164).

REFERENCES

- Ma Desi. Research on the Essence of Education Informatization: An Educational Perspective[D]. Shanghai: East China Normal University, 2007.
- [2] Dang Duo. Analysis of Services for Visually Impaired Readers in Public Libraries[J]. Chinese Information, 2022(5): 28-30.
- [3] Lu Qiujie. Practice and Reflection on Audio Description Services for Visually Impaired Individuals: A Case Study of Guangzhou Library[J]. Library World, 2020(6): 33-37.
- [4] Xu Jing. Research on APP Interface Design Based on the User Experience of Visually Impaired Individuals[J]. Shoe Technology & Design, 2023; 3(20): 128-130.
- [5] Yang Ruixue. Application of Speech Recognition Technology in News Interviews and Conversations[J]. Audio Engineering, 2024; 48(6): 59-61.
- [6] Tian Miao. Opening Up a New "Vision" for Services for Visually Impaired Groups: A Study on Innovative Reading Service Paths for Visually Impaired Groups in Public Libraries[J]. East South West North, 2022(15): 23-25.
- [7] Zamir M, Khattak D, Khan S, Rehman E. Smart Reader for Visually Impaired People Based on Optical Character Recognition[M]. 2020.
- [8] Goel A, Sehrawat A, Patil A, Chougule P, Khatavkar S. Raspberry Pi Based Reader for Blind People[C]. 2018.
- [9] Jadhav MS. Raspberry pi based reader for blind[J]. Int J Innov Emerg Res Eng, 2018; 5(1): 1639-1642.
- [10] Subbiah A. Camera based label reader for blind people[J]. Int J Chem Sci, 2016; 14(S3): 840-844.
- [11] Velmurugan D. A Smart reader for visually impaired people using Raspberry pi[J]. IJESC, 2016.
- [12] Bhargava A, Nath KV. Reading assistant for the visually impaired[J]. 2015; 5(2).
- [13] Shah PH. A portable prototype label reading system for blind[J]. 2015; 4(9).
- [14] Bisht S. Refreshable Braille Display using Raspberry Pi and Arduino[J]. 2016; 6(3).