Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 9
AIDML 2025

Research on the Application of Image-based Gesture Recognition Technology in Classical Industries

Yang Wang *

Aberdeen College of Data Science and Artificial Intelligence, South China Normal University, Foshan City, Guangdong Province, China

* Corresponding Author Email: 20223803037@m.scnu.edu.cn.

Abstract. In today's era of rapid technological development, the innovation of human-computer interaction technology is profoundly changing the way people live and work. Among them, gesture recognition technology, as a promising means of interaction, has gradually become an unstoppable trend in various classic industries. At present, gesture recognition technology is mainly divided into two categories: sensor-based data capture and vision-based non-contact perception. The former relies on high-precision sensors that can accurately capture subtle changes in movement; The latter uses advanced cameras and image processing technology to enable natural and contactless interactions. With the continuous progress of microelectronics technology and the wide application of deep learning technology, these two technologies have gradually shown a good trend of complementarity and integration, laying a solid foundation for the wide application of gesture recognition technology in many fields such as smart home, medical care, and transportation. However, the systematic comparison and evaluation of the performance boundaries, application scenarios, and user experience of these technologies is still insufficient. Through the analysis of sensor-based and vision technology, combined with practical application cases, this paper discusses its application potential and challenges in classic industries, and proposes corresponding future development directions.

Keywords: Gesture recognition; Image recognition; Visual perception; Deep learning.

1. Introduction

In today's era of rapid development of digitalization and intelligence, the innovation of human-computer interaction technology is profoundly changing the way people live and work. Gestures are an integral part of human daily life. Vision-based gesture recognition is a technique that combines complex perception with computer pattern recognition [1]. Although traditional interaction methods, such as keyboard and mouse, have played an important role in the development of computers, with the advancement of science and technology and people's growing demand for convenient and natural interaction, these traditional methods have gradually revealed their limitations. As an emerging, natural and intuitive means of interaction, gesture recognition technology is gradually emerging, and has shown broad application prospects in many classic industries. Traditional gesture recognition methods are mainly divided into sensor-based capture technology and vision-based non-contact perception technology [1]. The sensor approach relies on wearable devices to capture the user's gestures, while the vision approach relies on image processing and deep learning techniques to recognize gestures in real time.

However, although gesture recognition technology has made significant progress in theory and application, it still faces many challenges and problems. On the one hand, the systematic comparison and evaluation of the performance boundary, applicable scenarios and user experience of different gesture recognition technologies are relatively insufficient. This makes it difficult to accurately select and optimize the appropriate gesture recognition scheme according to specific needs in practical applications. In addition, the study has detected some potential effects on people due to advanced algorithms, such as any ethical issues that may arise in gesture recognition [2]. On the other hand, gesture recognition technology is also restricted by a variety of factors in practical application, such

as incomplete data and occlusion, and multiple degrees of freedom of the hand may lead to the occlusion of some areas or blurred posture when interacting with objects, thus affecting the accuracy of recognition. Environmental interference and lighting changes, complex lighting conditions in the real world, dynamic background noise and other environmental factors will also interfere with the effect of visual gesture recognition. In order to solve these problems, improvements need to be made in various aspects, such as the computing system needs to have good personalized adaptability, and the recognition system needs to have accurate recognition ability [3].

In view of this, this paper will deeply discuss the application potential and challenges of gesture recognition technology based on sensor and vision technology in classic industries, and put forward the corresponding future development directions. Through the analysis of related technologies and the study of practical application cases, it aims to provide useful references and suggestions for the further optimization and wide application of gesture recognition technology, and promote the effective implementation and innovative development of this technology in more fields, so as to better meet people's needs for natural and efficient human-computer interaction, further improve people's quality of life and work efficiency, and inject new impetus into the digital transformation and intelligent development of classic industries.

2. Comparison and Analysis of Typical Technologies

2.1. Sensor-based Gesture Recognition Technology

Sensor-based technology, as the name suggests, uses different sensors such as accelerometers and gyroscopes, while RGB cameras and infrared sensors are used to extract and recognize attributes from a set of gesture motion datasets, respectively, and a continuous gesture recognition (CHG) technique is proposed to continuously recognize hand movements [4]. Sensor technology relies on external sensors or wearable devices, such as bracelets, gloves, etc., to capture hand movements through sensors. The sensor technology has a high accuracy and response time, making it suitable for applications that require high accuracy. For example, in a car, the driver can use gestures to control the in-car entertainment system, navigation system, etc., by wearing gloves or other sensor devices. However, the disadvantages of sensor technology are also very obvious. First, the user must wear the device while this method requires the user to be physically connected to the computer, which may not be convenient enough for some users; Second, wearing the device may affect the user's natural movements, resulting in a less smooth interaction experience.

As a kind of sensor-based gesture recognition technology, the working principle of gesture recognition technology based on pulsed radio signals is as follows: the transmitter (Tx) emits pulsed radio waves through the antenna. The receiver (Rx) consists of an antenna, an amplifier, a low-pass filter, and a high-speed oscilloscope to receive signals reflected from the hand. When the hand is waved in the air, the amplitude and phase of the reflected waveform changes, and these changed waveforms can be used as features for analysis and identification by machine learning algorithms such as convolutional neural networks (CNNs) [5]. Convolutional neural networks can automatically extract key features from waveform data to accurately classify gesture actions.

Ultrasonic gesture recognition technology is also a major sensor-based gesture recognition technology, which uses speakers and microphones as ultrasonic input and output devices. The core principle is the use of ultrasound waves to generate Doppler shifts when they are reflected from moving objects, such as the hand [5]. During the gesture process, the system continuously samples the ultrasonic signal to obtain a series of data that changes over time. This data contains the unique characteristics of each gesture, such as the amplitude, frequency, and direction of the movement. In order to achieve accurate classification of unknown actions, this technology combines basic pattern recognition methods and supervised machine learning algorithms to comprehensively analyze and process the collected data. The basic pattern recognition method can perform the initial screening and classification of data, while the supervised machine learning algorithm can further improve the

accuracy and robustness of the classification. However, there are some limitations in the practical application of ultrasonic gesture recognition technology, such as adaptability in complex environments.

2.2. Vision-based Non-contact Perception Technology

Visual gesture recognition technology captures the user's hand movements through a camera or depth sensor and analyzes them through image processing algorithms. Unlike sensor technology, vision technology does not require the user to wear a device, has a better user experience, and is considered a common, suitable, and applicable technology because it provides contactless communication between humans and computers [6]. Especially in the smart home and medical fields, users can interact with the system without additional equipment, increasing the ease of use of the system.

Precise segmentation of the hand region is a key precursor to many vision-related tasks, and skin tone detection has become one of the most popular and widely used methods in the field of hand segmentation due to its unique advantages. Skin tone detection plays an integral role in many scenarios such as object classification, downgraded photo recovery, person motion tracking, video observation, human-computer interaction applications, facial recognition, and gesture recognition.

The implementation of skin tone detection relies on two core methods. The first is pixel-based skin detection, which independently compares each pixel in an image to a predefined skin tone model to determine whether it belongs to a skin area. This approach has the advantage of being fast and relatively simple to implement, but it also has obvious limitations, namely being susceptible to interference from factors such as complex backgrounds and lighting changes [7].

Another method is regional skin detection, which is no longer limited to the judgment of a single pixel, but broadens the field of view to the neighborhood area where the pixel is located, and comprehensively considers the multi-dimensional information such as the intensity and texture of the pixel in the area, and then determines whether the area is a skin area. To a certain extent, this method overcomes the influence of unfavorable factors such as lighting changes, and improves the accuracy and robustness of detection, but at the same time, it also increases the computational complexity and processing time [7].

However, skin tone detection methods are not flawless, and they come with a number of challenges. Dynamic changes in lighting conditions can cause skin tones to vary greatly in the image, which in turn can affect the accuracy of detection.

3. Application Case Analysis in Classic Industries

3.1. Smart Home

Many researchers have begun to explore the corresponding criteria and achievability of home automation systems under different home environments and user needs. Human-computer interaction (HCI), as a more interactive and resource-rich approach, can achieve efficient interaction with various devices, thereby ensuring the stable and smooth operation of home automation systems [8]. In the field of smart home, gesture recognition technology can realize the control of home devices. For example, users can use gestures to control the switching or adjustment of home devices such as lights, curtains, temperature, etc. Vision-based gesture recognition technology has a clear advantage in this scenario, as the user does not need to wear the device and can capture their movements directly through the camera. This contactless interaction makes the user experience more natural and convenient. However, complex background and lighting changes have a greater impact on the visual recognition system, which is also where the current smart home gesture recognition technology needs to be further optimized.

3.2. Medical Field

The application of gesture recognition technology in the medical field is of great significance, especially in the communication between patients and doctors, and gestures can convey emotions in a variety of ways, provide clues into the hearts of people, and help relieve anxiety and stress [9]. Gestures are a very effective alternative to communicating with deaf people, as they are able to express the other person's thoughts, even if verbal communication is not possible; For the elderly or patients with limited mobility, gesture recognition can be an effective way to interact with medical devices. For example, using gestures to control the lifting and lowering of hospital beds, adjusting the parameters of medical equipment, etc. In addition, vision sensor technology is widely used in medical gesture recognition systems to recognize patient gestures through image analysis and provide convenient contactless operation. However, there are also many challenges in the medical environment, such as hand occlusion, difficulty in identifying multiple skin objects, etc., which require further improvement of the technology.

3.3. Automotive Field

In automobiles, gesture recognition technology is used to control in-vehicle infotainment systems. For example, the driver can control the volume, switch songs, adjust the temperature of the air conditioner, and so on with simple gestures. Visual gesture recognition technology is especially suitable for operation during driving due to its contactless advantage, reducing the driver's distraction and operational complexity. Users can use simple gestures to control or interact with the device, allowing computers to understand human behavior, and the core technologies are gesture segmentation, gesture analysis, and gesture recognition [10]. However, challenges in the driving environment should not be overlooked, such as changes in lighting in the car, interference between the driver and other passengers, etc., which can affect the accuracy of gesture recognition.

4. Challenges and Prospects

4.1. Challenges

Gesture recognition has also exposed many problems that need to be solved in the current development. Incomplete data and occlusion: The high degree of freedom of the hand allows it to present a variety of poses and shapes, but this also leads to the problem of self-occlusion or blurred posture. For example, when the hand interacts with an object, parts of the hand may be obscured or difficult to capture clearly, which poses a significant challenge to vision-based gesture recognition systems [11]. Environmental interference and lighting changes: Changes in the real environment, such as different lighting conditions and dynamic background noise, can affect the effectiveness of visual gesture recognition. In the fields of smart home, medical care, and automotive, how to maintain stable recognition performance in complex environments is an important challenge for current technologies. Difficulty in finger recognition: Fingers, especially the fingertip parts, are highly similar in appearance, which makes it difficult for gesture recognition models to accurately distinguish between different finger movements. Improving the system's ability to recognize subtle differences is a difficult problem to be solved by current gesture recognition technology [11]. Individual differences and adaptability: Individual differences such as skin color, hand size, and movement habits of users will affect the accuracy of gesture recognition. Systems need to adapt to the characteristics of different users and provide personalized experiences, which requires technology to achieve efficient identification among diverse user groups.

4.2. Outlook

Many articles have assumed that the future direction of development is that gesture databases will be larger in terms of the number of gestures, the number of people in the database, and language coverage, and an important direction in the future of gesture recognition is to develop a database of multiple different sign languages, that is, a multilingual database that can be used for many future studies.

Gesture recognition can use a variety of characteristics, such as motion information, the position and direction of gestures. Feature extraction for dynamic gestures seems to be more challenging than static gestures. In addition to this, more attention should be paid to dynamic gesture recognition as well as continuous sign language. In addition, most existing research focuses on identifying isolated gestures, which are of limited use. The lack of continuous gestures may indicate the complexity of recognizing gestures and the inability of existing solutions to detect these gestures with acceptable accuracy. The future direction of gesture recognition may cover more areas. The future lies in the need to expand the current feature set to be able to recognize more gestures (e.g. those involving hands or facial expressions), and the possible future areas of this are the use of smart and wearable devices as data collection tools.

5. Conclusion

This paper first introduces two classic gesture recognition technologies, and gives examples to illustrate the specific applications of gesture recognition technology in some industries, and finally describes the current problems of gesture recognition technology and looks forward to the future. Sensors and visual recognition technologies are complementary in different application scenarios, with sensor-based technologies suitable for scenarios that require high accuracy, while vision-based technologies are suitable for interactions that do not require wearing devices. In the future, with the further development of deep learning and computer vision technology, the accuracy and robustness of gesture recognition will be significantly improved, especially in complex environments. In order to further promote the wide application of gesture recognition technology, researchers should pay attention to database construction, dynamic gesture recognition, personalized adaptability, etc., and develop more intelligent and efficient gesture recognition systems.

References

[1] M. Li, Z. Cui, Y. Yu, Computer vision-based hand gesture recognition for human-robot interaction: a review. Volume 10, pages 1581-1606 (2024).

[2] V. Chang, R.O. Eniola, L. Golightly, Q. A. Xu, An Exploration into Human-computer Interaction: Hand Gesture Recognition Management in a challenging Environment. Volume 4, article number 441 (2023).

[3] A. Ahmad, C. Migniot, A. Dipanda, Hand pose estimation and tracking in real and virtual interaction: A review. Image and Vision Computing. (2019).

[4] H.G. Prabhat, H.S. Chudga, S. Mukherjee, T. Dutta, K. Sharma, a Continuous Hand Gestures Recognition Technique for Human-Machine Interaction Using Accelerometer and Gyroscope Sensors. IEEE Sensors Journal (2016).

[5] F.A. Farid, N. Hashim, J. Abdullah, et al. A Structured and Methodological Review on Vision-Based Hand Gesture Recognition System. J. Imaging 2022, 8(6), 153; (2022).

[6] H. Kaur, J. Rani, A review: Study of various techniques of Hand gesture recognition. 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES) (2016).

[7] M. Oudah, A. Al-Najih, J. Chahl, Hand Gesture Recognition Based on Computer Vision: A Review of Techniques. J. Imaging 2020, 6(8), 73; (2020).

[8] B.I. Alabdullah, H. Ansar, N. Mudawi, et al. Smart Home Automation-Based Hand Gesture Recognition Using Feature Fusion and Recurrent Neural Network. Sensors 2023, 23(17), 7523; (2023).

[9] N. Mudawi, H. Ansar, A. Alazeb, et al. Innovative healthcare solutions: robust hand gesture recognition of daily life routines using 1D CNN. Volume 12 - 2024; (2024).

[10] S. Yu, Thoughts on Artificial Intelligence in the Automotive Field. China Strategic Emerging Industries, (32), 125. (2018).

[11] K. GAO, H. Zhang, X. Liu, et al. Challenges and solutions for vision-based hand gesture interpretation: A review. Computer Vision and Image Understanding, Volume 248, November 2024, 104095 (2024).