Research Article

Design and Application of Interactive Algorithm for Advertising Media Screen Based on Smart Sensor

Si Chen ⓘ, Dandan Cheng, and Quan Zhou

Department of Information Technology, Zhengzhou Vocational College of Finance and Taxation, Zhengzhou 450000, China

Correspondence should be addressed to Si Chen; 12006001@zzcsjr.edu.cn

Touch is one of the most important human senses. With the popularization of touch-screen mobile phones, tablet computers, and other devices, touch-screen interactive technology has become a norm in people's daily lives, and advertisements that were once dominated by vision and hearing have added an interactive experience in the dimension of touch. Traditional advertising media screens can only complete simple information dissemination functions and cannot interact with users in a two-way manner. They can only receive information one-way and passively and lack interactivity. Touch-screen interactive advertising forms a good interaction with the target audience, thereby disseminating advertising information to achieve the purpose of promotion or brand image building. This paper designs a set of advertising media screen interaction systems based on smart sensors, including a gesture interaction module, a remote interaction module, and a touch interaction module. The gesture interaction module can recognize 5 static gestures and send gesture commands to control the advertising media screen. The remote interaction module can remotely control the advertising media screen, and the touch interaction module can control the advertising media screen through the touch screen. According to the functional requirements, the overall design of software and hardware is given, and the technical background of each module of the software is introduced. Next, the depth image-based gesture recognition method is studied. The number of fingers and the center distance feature are fused as feature vectors, and the weighted template matching method is used to classify and recognize gestures. Finally, the design and implementation of the interactive system are introduced.

1. Introduction

The sense of touch is not only an organ for human beings to feel the surrounding environment, but can also express intimacy, kindness, tenderness, and caring. Emotional communication among friends, lovers, and relatives all relies on the sense of touch. It is a sick society. It is no longer possible to recognize the reality of things by sight and hearing, and only the sense of touch can approach reality. Therefore, human beings have an innate trust in tactile cognition, and tactile sensation is more realistic than any sensation [1–4]. The machine originally invented by mankind mainly uses vision and hearing for human-computer interaction. With the development of tactile technology and a decrease in cost, devices that can provide tactile input and output have become more and more popular, and touch screen mobile phones are the most popular devices.

Although the current technology of using touch screens for human-machine dialogue is still a bit naive, at least people can get rid of the cold mouse and buttons, petting the machine like petting an animal and getting feedback. The upgrade of this experience is a milestone [5–10]. They are private and usually have only one user, emphasizing a personalized experience; public touch-screen media mainly refers to touch-screen devices set up in public places, public information systems, etc., which are public, usually have various users, and emphasize a popular experience [11–16]. The advertising media screen is shown in Figure 1.

Advertising creative personnel use touch screen interactive technologies, such as mobile photography, QR code recognition, APP applications, positioning systems, voice recognition systems, and other new technologies, combined with some traditional media, such as outdoor billboards, TV advertisements, magazine advertisements, etc. develop

FIGURE 1: Advertising media screen.

advertising campaigns. Advertisers design advertising information into the interactive link, and advertising audiences use touch screens to interact with people and receive advertising information during the process of participation. These can be called "touch-screen interactive advertising." [17–21] The media in touch-screen interactive advertising is a "touch-screen" type of media, the experience mode is "interactive," and advertising is its goal. The purpose of touch-screen interactive advertising is to form a good interaction with the target audience, thereby disseminating advertising information to achieve the purpose of promotion or brand image building [22–24].

The initial gesture recognition is to use the data glove to directly obtain information about the human hand area. The shape of the data glove is similar to an ordinary glove; the difference is that the sensor is installed on the surface. S. Sidney Fels and Geoffrey E. Hinton used VPL data gloves to implement a gesture-speech system that can recognize 203 vocabulary words. Ji-Hwan Kim and others developed a three-dimensional gesture tracking and recognition system using KHU-1 data gloves, which transmit gesture signals to a PC via Bluetooth, which can recognize simple gestures such as scissors, rocks, and cloth. The Data Glove 14 ultra data glove developed by 5DT Company can measure the curvature of the finger. The material is made of stretchable synthetic elastic fiber. 14 optical fiber sensors provide an 8-Bit open bending feature to adapt to the use of different palm sizes. Based on the Data Glove14 ultra data glove, Li Dongjie and others proposed a recognition method combining BP neural network and the PSO algorithm, which improved the recognition rate [15, 25, 26].

Compared with a gesture recognition system based on data gloves, a vision-based gesture recognition system enables users to interact with humans more naturally. Henrik Birk et al. used principal component analysis to analyze 1000 gesture pictures and selected feature parameters to realize gesture recognition for 1500 pictures, with a recognition rate of 99%. Lee et al. used entropy analysis to segment and extract the gesture region from the video stream for gesture recognition. The background in the video stream is more complex. The system can recognize 6 kinds of gestures, and the recognition rate can reach 95%. Bobick and Wilson et al. applied a state-based gesture recognition method. The gesture is equivalent to a trajectory in space. The trajectory curve is divided into states, and the gesture is represented as a set of continuous sequences. At the beginning of recognition, the image to be recognized is compared with the prototype of the trajectory curve state, and the result of gesture recognition is obtained. Shen Qing et al. proposed a template-based gesture recognition framework that can detect and segment actions in real-time during video-based human-computer interaction. This method can process online video sequences in real-time. Yang Bo et al. proposed a gesture recognition algorithm based on the spatial distribution characteristics of gestures. The gesture location algorithm uses the "search window" to filter the current skin color area, and the gesture segmentation algorithm uses a gesture segmentation algorithm based on the brightness Gaussian model to select spatial relative density features and knuckles. The relative distance between the features is used as a feature vector, and the similarity is calculated by using the feature vector to achieve the purpose of recognizing gestures. The recognition rate can reach 98% [27, 28].

Traditional advertising media screens can only complete simple information dissemination functions and cannot interact with users in a two-way manner. They can only receive information one way and passively and lack interactivity. Touch-screen interactive advertising forms a good interaction with the target audience, thereby disseminating advertising information to achieve the purpose of promotion or brand image building.

2. Interactive Algorithm

2.1. Gesture Recognition. Gesture, as a means of interaction that conforms to a human's daily interaction habits, is simple, direct, and natural. The research and application of gestures has attracted great attention from researchers at home and abroad. Compared with traditional camera technology, TOF sensor camera technology can obtain accurate depth information of objects, enrich image feature information, and have a huge boost in gesture recognition.

In the process of image acquisition, transmission, and transformation, it will inevitably be disturbed, and noise will affect the image quality. Therefore, it is necessary to preprocess the collected hand region image, remove the noise part of the image, and enhance the effective image. The preprocessed image has a better effect and is easier to recognize than the original image. At present, the commonly used image filtering includes median filtering, mean filtering, Gaussian filtering, etc. Several classic image filtering methods are introduced below.

2.1.1. Mean Filtering. The idea of mean filtering is to first use the EPC660 TOF camera to shoot multiple consecutive depth image sequences, then randomly select a certain number of images from the shooting sequence, and finally sum the depth value of each pixel to find the average value to

reduce noise. This kind of average filtering is suitable for the situation when the measured object has a small displacement or is still. When the measured object is in a moving state, if the average filter is used, the image will be blurred because the pixels at the same point in different images are parallel. It is not the same position of the object being measured. When the noise in the image is too concentrated in a certain area, the average filter is used, and the filtering effect is not ideal. In addition, the continuous acquisition of multiple images to calculate the average will also cause the real-time performance to decrease. When the application scenario requires high real-time performance, using the method of mean filtering is not applicable; when the number of selected images is more or less, it will also affect the filtering effect.

2.1.2. Median Filtering.

Median filtering is a local nonlinear smoothing filtering method, proposed by Turkey in 1971. The basic idea of median filtering is that for any point in the image, we select pixels in its neighborhood, sort these pixels according to the gray level, and find the median value in the pixel sequence as the current pixel value. The algorithm of median filtering is as follows:

$$F(x, y) = \text{mid}\{f(x-k, y-l), k, l \in W\}, \tag{1}$$

where F and f, respectively, represent the pixels after median filtering and the pixels before filtering, W is the size of the filtering window, and k, l is the length and width of the window.

2.1.3. Gaussian Filtering.

We analyze the image from the signal point of view. The noise of the image belongs to the high-frequency part, and a low-pass filter can be used to reduce the noise. Gaussian filtering is a typical linear smoothing filtering method that is widely used in image noise reduction. The implementation method is that each pixel in the image is scanned through a template. The template will determine a neighborhood pixel area, and the weighted average of the pixels in this area is used to replace the value of the center pixel of the template. The template can be obtained in the following ways:

$$h(x, y) = \frac{1}{2\pi\sigma^2} e^{-\left(x^2+y^2/2\sigma^2\right)}, \tag{2}$$

where x and y are the point coordinates, and σ is the standard deviation. The Gaussian filter template can be obtained by discretizing the Gaussian function. We bring the position coordinates into the Gaussian function, and what we get is the template coefficient. The results of the template with a size of 3×3 and $\sigma = 0.8$ are normalized and the following results are obtained:

$$h = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \tag{3}$$

Touch screen technology first moved from the laboratory to civilian equipment, such as touch screen mobile phones and tablet computers, allowing ordinary people to enjoy the convenience of finger control, and then, due to the popularization of touch screen technology, various types of touch screen media were derived, which became advertisements. The carrier of the message according to the template obtained from the above derivation and calculation results, the depth image is weighted and denoised. The formula is as follows:

$$F(x, y) = \sum_{m=0}^{L} \sum_{n=0}^{L} f(x-m, j-n) \times h(m, n), \tag{4}$$

where F is the pixel value after filtering, L is the size of the template window, f is the pixel value before filtering, and h is the Gaussian template.

Touch screen interactive design is a new design thinking based on touch screen media. Simply put, when the audience uses touch-screen media such as touch-screen mobile phones and tablet computers, they can "design thinking of direct two-way communication with the system and other recipients to a certain extent through a variety of input and output methods." The template matching method is currently the most widely used classification method in static gesture recognition. The template is a vector designed to detect certain regional features. As shown in Figure 2, its basic principle is to obtain a standard template for each class to be identified as the identification standard. By comparing the degree of similarity between the class to be identified and the standard template, the highest similarity is the identification result. The standards of similarity mainly include the following: Minkowski distance, Euclidean distance, Manhattan distance, and cosine angle. Among them, the following formula is for two p-dimensional array objects:

$$\begin{aligned} a &= \left(x_{i1}, x_{i2}, ..., x_{ip}\right), \\ b &= \left(x_{j1}, x_{j2}, ..., x_{jp}\right). \end{aligned} \tag{5}$$

The calculation formula of Manhattan distance is as follows:

$$D(a, b) = \left|x_{i1} - x_{j1}\right| + \left|x_{i2} - x_{j2}\right| + \cdots + \left|x_{ip} - x_{jp}\right|. \tag{6}$$

The formula for calculating Euclidean distance is as follows:

$$D(a, b) = \sqrt{\left(\left|x_{i1} - x_{j1}\right|^2 + \left|x_{i2} - x_{j2}\right|^2 + \cdots + \left|x_{ip} - x_{jp}\right|^2\right)}. \tag{7}$$

For more generalization, it becomes

$$D(a, b) = \left(\left|x_{i1} - x_{j1}\right|^q + \left|x_{i2} - x_{j2}\right|^q + \cdots + \left|x_{ip} - x_{jp}\right|^q\right)^{1/q}. \tag{8}$$

When $q = 1$, it is Manhattan distance, and when $q = 2$, it is Euclidean distance. The above three formulas meet the following conditions:
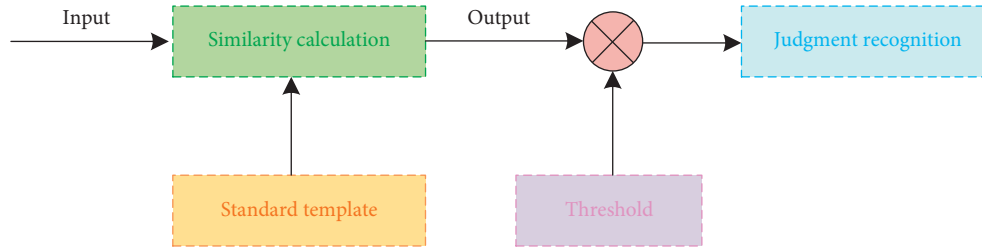
Figure 2: Template matching.

$D(a, b) \geq 0$: distance is positive,

$D(a, b) = 0$: distance is 0,

$D(a, b) = D(b, a)$: distance is symmetric,   (9)

$D(a, b) \leq D(a, k) + D(k, b)$: distance less than sum.

The calculation formula of the included angle cosine is as follows:

$$\cos(\theta) = \frac{a \cdot b}{|a||b|}. \qquad (10)$$

2.2. Related Technologies of Remote Interaction Module.
MQTT (message queuing telemetry transport), as an instant messaging protocol, was released by IBM in 1999. The protocol supports all platforms and is a "lightweight" communication protocol based on the publish/subscribe mode. The protocol is based on the TCP/IP protocol and can provide real-time reliability for remote devices under limited bandwidth. Messaging service. Because of its low overhead and low bandwidth occupancy characteristics, the current MQTT protocol has a wide range of applications in the Internet of Things, mobile applications, and small devices. MQTT is a client-server message publish/subscribe transmission protocol. Publisher (Publish), subscriber (Subscribe), and broker (Broker) are the three different identities in the protocol. The message broker is completed by the server, and the client is responsible for publishing and subscribing to messages.

We can divide touch-screen media into two categories based on different audiences: personal touch-screen media and public-place touch-screen media. Personal touch-screen media mainly refers to touch-screen media devices such as mobile phones and tablets that are easy to carry around. Topic and Payload are messages transmitted by the MQTT protocol. A topic is the type of message. After the client subscribes, it will receive the content of the message. The payload is the content of the message, which is used by the subscriber. As shown in Figure 3, a device that uses the MQTT protocol must consist of a client and a server. The client can perform operations such as publishing messages, subscribing to messages, unsubscribing to messages, and disconnecting.

The message header of all MQTT messages must contain a fixed message header. From the perspective of saving overhead and reducing network traffic consumption, the message header is designed with a fixed size of 2

Bytes. The special message contains a variable header and a payload.

The way in which the client and the server ensure that the message is received is determined by QoS. The client can choose topics of any QoS level to subscribe to. When topics with different QoS levels exist at the same time, the client will choose the one with the highest QoS level to receive. An example is as follows: a message is published at the QOS2 level. At this time, there are two clients that need to subscribe to this message. Client A subscribes to the QoS0 level, so this message will be sent to client A according to QoS0, and client B uses QoS2 to subscribe to the same message. At this time, when client A receives the same message, it becomes the QoS2 standard. When publishing a message, this field is valid. When RETAIN is 1, it means that the sent message needs to be kept persistent, even if the server restarts, it will be kept, not only to the current subscribers but also new subscribers who subscribe to this topic in the future will be pushed immediately; RETAIN is 0. At the time, this message is only pushed to the current subscriber. When the server or client tries to resend SUBSCRIBE, PUBLISH, and PUBREL messages, the DUP flag is set. If the QoS level of the message is greater than 0, a message confirmation is required. At this time, the DUP flag can be set, and the message ID will also be included in the variable header. The meaning of this field is that the remaining bytes of the current message include variable headers and payloads, and Byte2 is exclusively occupied by this field. The maximum value of a single byte: 01111111, hexadecimal: $0 \times 7F$, and decimal is 127.

2.3. Related Technologies of Touch Screen Interactive Module.
Qt is the current mainstream tool for embedded graphical interface development. It is a C++-based, cross-platform graphical user interface program framework. The Qt framework was released by the Qt Company in 1991 and has been developed to Qt 5.10. The characteristics of Qt are as follows:

(1) The mechanism of signals and slots ensures that the communication between classes within the program is safe and effective and avoids memory field pointers and object life cycle problems caused by shared pointers

(2) Qt supports Windows, Linux, Android, Mac OS, and other operating systems, with excellent cross-platform features

Figure 3: The structure of MQTT.

(3) Provide a rich class library, including file processing, IO interface processing, expression processing, etc.

(4) Support 2D and 3D graphics rendering

(5) Fully object-oriented development platform, high code reusability, encapsulating a large number of class modules, developers can easily extend and use these modules

(6) Support QML, users can easily describe the user interface by using QML

(7) A large number of development documents are for developers' reference

An important prerequisite for gesture recognition is to segment the human hand area from the complex background environment, and then extract the features of the gesture for the next step of recognition. Qt's object-to-object communication adopts its signals and slots (signals and slots) mechanism. Signals and slots are an alternative to the callback technique. Before the signal and slot mechanism work, a connection must be established. After a certain event triggers the signal, the slot that has established a connection with the signal will respond, that is, perform the operation of the slot function. The correspondence between signals and slots can be one-to-one, one-to-many, many-to-many, and many-to-one. When a signal connects multiple slots, the execution order of the slot functions is random. The signaling object does not need to know which slot an object needs to receive the signal it sends, but only needs to send the signal when appropriate, and does not pay attention to whether the transmission is successful. Similarly, the slot of the object does not know which signals are associated with itself. This mechanism greatly reduces the coupling of the Qt object.

The connection between signals and slots is mainly divided into queue connection, automatic connection, direct connection, and blocking queue connection. A queue connection means that the object that triggers the signal and the object that receives the signal are not in the same thread and return immediately after the signal is sent, until the thread where the slot is located gets control and starts the event loop. Automatic connection means that when the signal is triggered, the mechanism will select the direct connection and the queue connection according to whether the object triggering the signal and the object receiving the signal are in the same thread. Direct connection means that the object that triggers the signal and the object that receives the signal are in the same thread, and the slot function will be called immediately after the signal is triggered. Blocking queue connection means that the thread that triggers the signal is in a blocked state after the signal is triggered and then resumes from the blocked state to the running state after the slot function returns.

3. Gesture Recognition Based on Smart Sensors

These methods are easily affected by environmental factors, resulting in difficult segmentation and poor results. In this paper, the EPC660 TOF sensor is used to collect the depth image, and the hand region segmentation and feature extraction are completed based on the depth image. The gesture area segmentation method proposed in this paper is based on the premise that the hand area is the area closest to the EPC660 TOF sensor. Because during human-computer interaction, the user will naturally stretch his hand to the front of the body to make gestures. At this time, the hand area is generally in the area closest to the sensor. Based on the above, the segmentation of the hand area becomes easier. That is, the set of pixels closest to the sensor is the approximate area of the hand. In this paper, an automatic threshold segmentation algorithm based on a depth histogram is used to segment the hand region.

The gesture features include shape and contour. The main methods of gesture region segmentation based on two-dimensional images are based on skin color models, contour models, and wearing gloves that are different from the background color. The histogram is a two-dimensional statistical report graph. The horizontal and vertical coordinates reflect the measurement of a statistical sample and a certain attribute corresponding to the sample, and they are widely used in computer vision analysis. A depth image is composed of pixels with different depth values, and the depth distribution in the image is an important feature of the image. The depth histogram describes the distribution of the depth values in the image in the distance interval and can intuitively show the frequency and aggregation of each depth in the image. With this information, the segmentation threshold can be determined. As shown in Figure 4, the left image is the depth histogram corresponding to the right image.

Observing Figure 4, it can be found that the pixels of the depth image are mainly concentrated in the three distance intervals, most of the pixels are concentrated in the interval on the right. These pixels are part of the background environment farthest from the camera. This part is useless information. We discard it directly; the pixels in the middle interval correspond to the pixels of the human torso, which also needs to be discarded; the left interval is the set of pixels closest to the camera, this part of pixels is valid information and needs further processing. It is determined through experiments to sort the leftmost interval from small to large and find the interval with the least number of pixels. The distance range of this interval is the optimal segmentation threshold point F, and this distance is used as the dynamic threshold for gesture region segmentation.
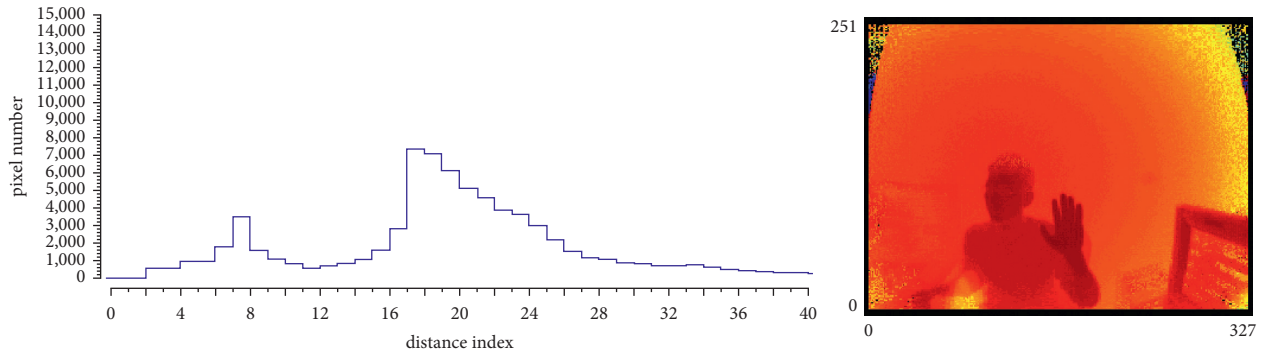
FIGURE 4: Depth histogram.

Although the EPC660 TOF sensor used in this paper has the characteristics of high frame rate and high measurement accuracy, according to error analysis, the image will inevitably be interfered with and generated during the process of acquisition, transmission, and transformation, which will reduce the image quality. Therefore, it is necessary to preprocess the collected hand region image, remove the noise part of the image, and enhance the effective image. The preprocessed image has a better effect and is easier to recognize than the original image. This paper has carried out filtering and denoising processing on the image. Currently, the commonly used image filtering includes median filtering, mean filtering, Gaussian filtering, and so on.

The most important prerequisite for gesture recognition is that the gesture features are accurate, which can uniquely characterize a certain gesture and distinguish it from other gestures. If the selected feature parameters are not appropriate, the recognition results may also be inconsistent for the same gesture. Therefore, the selection of gesture features directly affects the result of gesture recognition. The selection of gesture features should follow the following principles:

(1) *Uniqueness.* The gesture feature can only uniquely determine one gesture under any conditions;

(2) *Completeness.* All features of a gesture should be represented by the features of this gesture;

(3) *Geometric invariance.* The external conditions of the human body, such as height, weight, and skin color, should not affect gesture characteristics. Generally, the commonly used gesture features include the number of fingers, the angle between the fingers, the overall outline of the hand, etc. If the recognition of the gesture is more complicated, multiple features need to be extracted at the same time to form a feature vector to express the gesture.

Since the difference between the five predefined digital gestures in this article is the number of fingers extended, the gesture feature can be selected as the number of fingers. For this feature, the main task is to identify the number of fingertips. At the same time, a single feature cannot guarantee the accuracy of recognition, and multiple features need to be searched. The predicted amplitude is shown in Figure 5.

4. Hardware System

In this part, a gesture recognition module is designed. According to the result of gesture recognition, the advertising media screen is controlled according to the predefined gesture meaning. Figure 6 shows the gesture recognition module architecture. The lower computer and the upper computer use TCP communication. The lower computer server runs under the BeagleBone Black single-board Linux system of the EPC660 TOF sensor camera module and is responsible for collecting image information; the upper computer client runs on the PC Windows system, development of a visual graphical interface based on Qt, responsible for displaying and processing the acquired image information, and performing gesture recognition, and finally converting the recognition result into a control message and sending it to the playback module. The playback module executes the command and returns the execution result. The host computer uses UDP unicast communication with the playback module.

Since advertising media screens are generally located on the outer walls of high-rise buildings, and multiple advertising media screens may be scattered in different locations and far apart, in order to facilitate users to configure equipment, it is necessary to design a set of remote interaction schemes to realize remote control of advertising media screens. Control and monitoring based on actual engineering requirements, this paper develops a remote interactive client based on the MQTT protocol and cooperates with remote server instructions to realize remote interactive functions.

4.1. Remote Interaction Design Based on MQTT Protocol. The remote interaction module based on the MQTT protocol designed in this article consists of three parts, including the cloud part, the MQTT client part, and the MQTT server part. The cloud is responsible for interacting with the user's web page, sending the user's instructions through the MQTT server to publish a "Publish" message to the advertising media screen MQTT client, and the MQTT client is responsible for replying to the "Publish" message to the MQTT server. Messages use the MQTT protocol for downlink and the http protocol for messages. The client runs on the Linux platform of the advertising media screen controller, and the
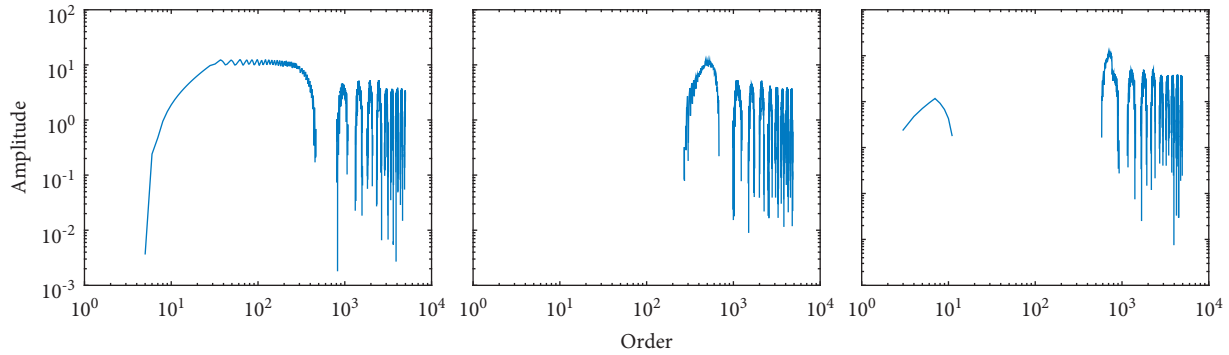
FIGURE 5: Predicted amplitude.

FIGURE 6: The architecture of the gesture recognition module.

MQTT server uses the server provided by the Alibaba Cloud Internet of Things suite.

The advertising media screen has the following application scenarios: multiple advertising media screens are set up on multiple building curtain walls to synchronously play the same program. At the same time, because the output resolution of a single advertising media screen controller is limited by the network bandwidth, a large-resolution advertising media giant screen needs to be composed of multiple advertising media screens, which are controlled separately by multiple advertising media screen controllers. The advertising media screen is responsible for playing a part of the video. In this way, multiple advertising media screens must be played synchronously, with a mutual error of no more than 5 frames, in order to play a complete screen at the same time. Figure 7 shows the evaluated value.

This paper designs a set of multiscreen synchronization interactive solutions based on GPS timing, which can ensure time synchronization through GPS timing when multiple multimedia block advertising media screen controllers operate independently and are deployed over a long distance. On this basis, the frame control algorithm is used to achieve the effect of synchronized playback of multiple advertising media screens. The entire method includes four modules, GPS timing module, program list analysis module, frame control module, and video playback module. The GPS timing module is responsible for calibrating the time of the multimedia large-screen controller; after the GPS timing module starts running, we open the GPS data serial port to establish a data interaction channel with the main process of the multimedia large-screen controller; we obtain GPGGA data from the GPS data serial port. GPGGA is GPS. The data output format sentence is the main data of a frame of GPS positioning; we parse the world time from GPGGA data,

store it in the buffer, and send the buffer data to the main process of the multimedia large-screen controller for time calibration; and get it with a delay of ten minutes. Next time we get GPGGA data, we will continue to time service. The playbill parsing module is responsible for querying the XML file of the total playbill and generating a list of play files. The total playbill XML file contains multiple play items, and each play item contains information such as start play time, stop play time, play date or week set, play mode, and play file path. y versus x is shown in Figure 8.

After the module starts to run, we load the total playbill XML file and start to generate the play file list XML file. First, we traverse the XML file of the total playbill, we compare the current time T with the stop play time T1 of the current traversed play item S, if T1 is greater than T, then we determine the play mode of S. If T1 is less than T, it means that the current time is not in the play time period of S. Within the range, we traverse the next play item directly. There are four play modes: weekday mode, holiday mode, week mode, and loop mode. If the playing mode of S is holiday mode or weekday mode, it is judged whether the current date is the same as the date of S, and if they are the same, S is written into the play file list XML file. If they are not the same, we traverse the next play item; if the play mode of S is not holiday mode or weekday mode, we determine whether the play mode of S is week mode; if the play mode of S is week mode, we calculate the week of the current date, then judge whether it is in the week set of S; if it is, we write S into the play file list XML file, otherwise, we traverse the next play item. If the play mode of S is not week mode, then we judge whether the play mode of S is loop mode; if the play mode of S is loop mode, we write S into the play file list XML file, otherwise traverse to the next play item. The frame control module is responsible for controlling the speed at which
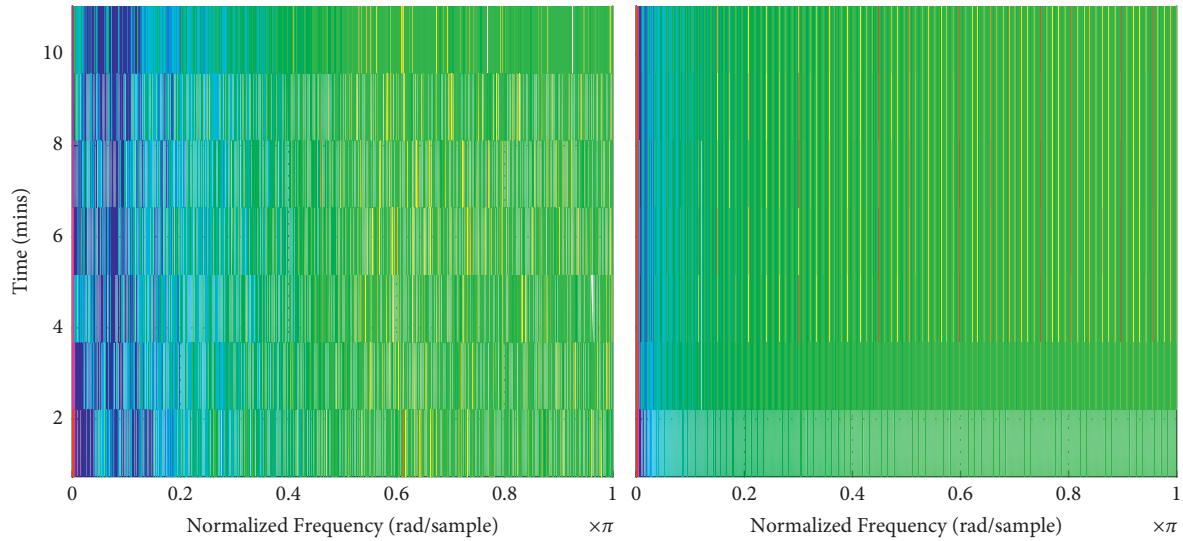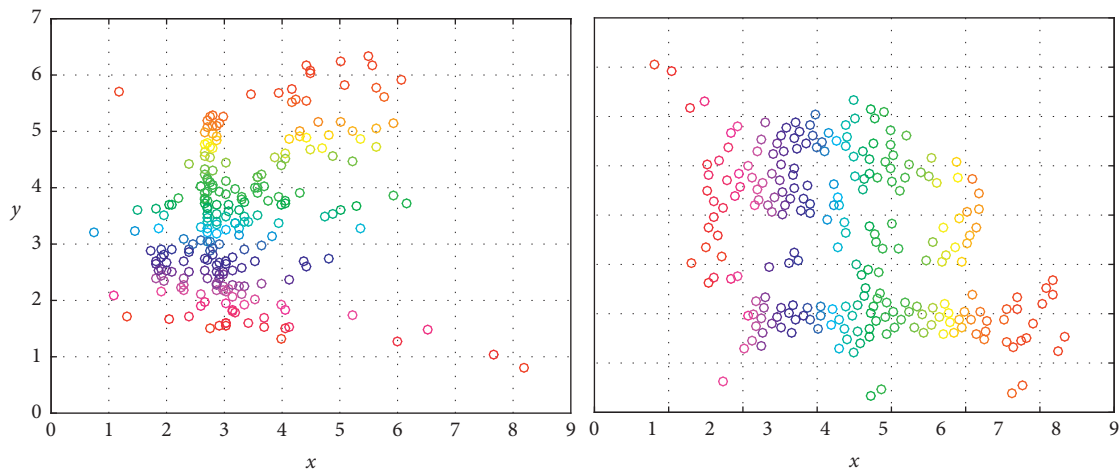
FIGURE 7: Evaluated value.

FIGURE 8: y versus x.

video data is transmitted to the video playback module. First, we parse the play file list XML file generated by the playbill parsing module, and then read the file information in sequence according to the play file list sequence. The file information includes video length, frame rate, and total frame number; if the start playing time stamp configuration file does not exist, it will be read accordingly. The video length in the file information calculates the start playing time of the next video, and after storing it in the start playing timestamp configuration file, reads one frame of data; if it exists, reads the start playing of the current video from the start playing timestamp configuration file, timed by the timer to start playing time; then we read a frame of data to judge whether it is the end frame. If it is the end frame, we judge whether the currently playing file is the last playing file, if yes, we exit the frame control module, if not, the last file to be played will continue to read the information of the next file and play the next file. If it is not the end frame, we adjust the timer timing

T through the PID algorithm. First, we calculate the expected number of frames by multiplying the frame rate by the playing time; then make the difference with the actual number of frames fed back by the video playback module to obtain the frame number deviation, and calculate the adjusted value through the derivative term D, the proportional term P, and the integral term I T value; after the precise time T is timed by the timer, the read video frame is sent to the video playback module to continue to read the next frame of data. The video playback module is responsible for constructing a video pipeline, decoding, and playing video data. After the video playback module starts to run, first we build a video pipeline, open up a buffer for storing video data, and wait for the data to arrive; after receiving the data, we determine whether it is video data, if not, we continue to wait; if it is video data, the data will be sent to the video pipeline for decoding and playback, and then continues to wait for the data. Figure 9 shows the successful recognition gesture.
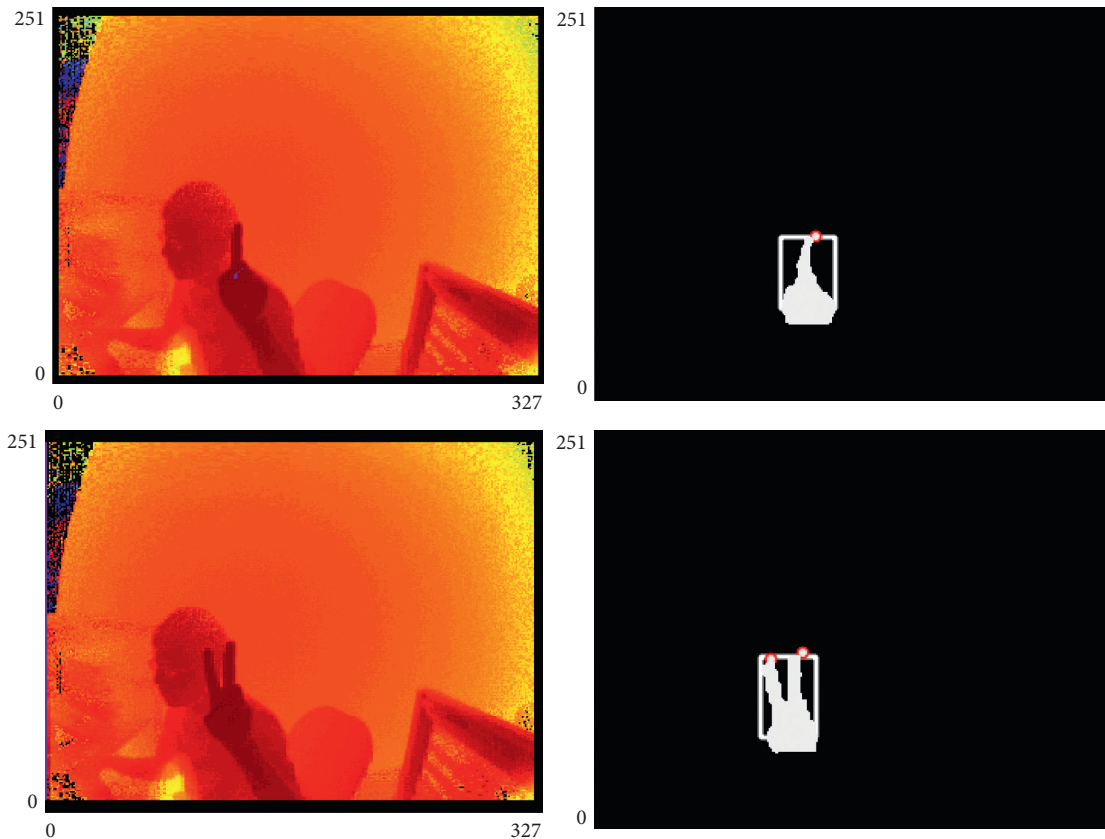
FIGURE 9: Successful recognition gesture.

5. Conclusion

This paper combines the actual engineering needs to meet the interactive needs of different users. For the gesture recognition function, the sensor module is used to conduct in-depth research on the principle of depth image acquisition, and the gesture recognition is realized by using the depth image. For the touch screen interaction function, the touch screen program design based on the Qt framework is adopted.

The advertising media screen interactive software designed in this paper has been successfully used in many practical projects, and the operation effect is good, but there are some shortcomings, such as only 5 simple static gestures can be recognized, and there is no support for complex dynamic gesture recognition. The design of the software interface is not beautiful enough and needs to be improved in the next step.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest with this study.

Acknowledgments

References

[1] Z.-J. Pang, R.-Z. Liu, Z.-Y. Meng, Y. Zhang, Y. Yu, and T. Lu, "On reinforcement learning for full-length game of Star-Craft," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 4691–4698, 2019.

[2] O. Schmidt, A. Gambhir, I. Staffell, A. Hawkes, J. Nelson, and S. Few, "Future cost and performance of water electrolysis: An expert elicitation study," *International Journal of Hydrogen Energy*, vol. 42, no. 52, pp. 30470–30492, 2017.

[3] L. C. Astfalck, E. J. Cripps, J. P. Gosling, M. R. Hodkiewicz, and I. A. Milne, "Expert elicitation of directional metocean parameters," *Ocean Engineering*, vol. 161, pp. 268–276, 2018.

[4] H. Landquist, "Expert elicitation for deriving input data for probabilistic risk assessment of shipwrecks," *Marine Pollution Bulletin*, vol. 125, no. 1-2, pp. 399–415, 2017.

[5] M. Jamshidi and F. Ghanati, "Taxanes content and cytotoxicity of hazel cells extract after elicitation with silver nanoparticles," *Plant Physiology and Biochemistry*, vol. 110, pp. 178–184, 2017.

[6] J. Yu and P. Lu, "Learning traffic signal phase and timing information from low-sampling rate taxi GPS trajectories," *Knowledge-Based Systems*, vol. 110, no. 1, pp. 275–292, 2016.

[7] K. P. Wijayaratna, V. V. Dixit, and L. Denant-Boemont, "An experimental study of the Online Information Paradox: Does

en-route information improve road network performance?" *PLoS ONE*, vol. 12, no. 9, pp. 184–191, 2017.

[8] Z. Wang, H. Ren, Q. Shen, W. Sui, and X. Zhang, "Seismic performance evaluation of a steel tubular bridge pier in a five-span continuous girder bridge system," *Structures*, vol. 31, no. 1, pp. 909–920, 2021.

[9] S. Nakayama and J. Takayama, "Traffic network equilibrium model for uncertain demands," in *Proceedings of the 82nd Transportation Research Board Annual Meeting*, Washington, DC, USA, 2021.

[10] D. Matekalo, M. Skorić, T. Nikolić et al., "Organ-specific and genotype-dependent constitutive biosynthesis of secoiridoid glucosides in *Centaurium erythraea* Rafn, and its elicitation with methyl jasmonate," *Phytochemistry*, vol. 155, pp. 69–82, 2018.

[11] J. A. S. Shantz, J. R. S. Leiter, T. Gottschalk, and P. B. MacDonald, "The internal validity of arthroscopic simulators and their effectiveness in arthroscopic education," *Knee Surgery, Sports Traumatology, Arthroscopy: Official Journal of the ESSKA*, vol. 22, no. 1, pp. 33–40, 2014.

[12] M. A. Boussadi, T. Tixier, A. Landrault, and J.-P. Derutin, "HNCP: A many-core microprocessor ASIC approach dedicated to embedded image processing applications," *Microprocessors and Microsystems*, vol. 47, pp. 333–346, 2016.

[13] M. Király and I. Pulido, "Real-time nanopositioning for security holograms," *Photonics Views*, vol. 16, no. 1, pp. 64–66, 2019.

[14] M. Conti and A. Passarella, "The Internet of people: A human and data-centric paradigm for the next generation Internet," *Computer Communications*, vol. 131, pp. 51–65, 2018.

[15] M. Grossi, "A sensor-centric survey on the development of smartphone measurement and sensing systems," *Measurement*, vol. 135, pp. 572–592, 2019.

[16] A. Cuzzocrea, "Intelligent algorithms for data-centric sensor networks," *Journal of Network and Computer Applications*, vol. 35, no. 4, pp. 1175-1176, 2022.

[17] A. Ignatov, "Real-time human activity recognition from accelerometer data using Convolutional Neural Networks," *Applied Soft Computing*, vol. 62, pp. 915–922, 2018.

[18] Y. Lu, C. Zhang, B.-Y. Zhou, X.-P. Gao, and Z. Lv, "A dual model approach to EOG-based human activity recognition," *Biomedical Signal Processing and Control*, vol. 45, pp. 50–57, 2018.

[19] S. Feng and M. F. Duarte, "Few-shot learning-based human activity recognition," *Expert Systems with Applications*, vol. 138, p. 112782, 2019.

[20] M. M. Hassan, M. Z. Uddin, A. Mohamed, and A. Almogren, "A robust human activity recognition system using smartphone sensors and deep learning," *Future Generation Computer Systems*, vol. 81, pp. 307–313, 2018.

[21] Y. Chen, Y. Ma, X. Mao, and Q. Li, "Multi-task learning for abstractive and extractive summarization," *Data Science and Engineering*, vol. 4, no. 1, pp. 14–23, 2019.

[22] P. Zhou and Z. Jiang, "Self-organizing map neural network (SOM) downscaling method to simulate daily precipitation in the Yangtze and Huaihe River Basin," *Climatic and Environmental Research*, vol. 21, no. 5, pp. 512–524, 2016.

[23] H. Beitollahi, Z. Dourandish, S. Tajik, M. R. Ganjali, P. Norouzi, and F. Faridbod, "Application of graphite screen printed electrode modified with dysprosium tungstate nanoparticles in voltammetric determination of epinephrine in the presence of acetylcholine," *Journal of Rare Earths*, vol. 36, no. 7, pp. 750–757, 2018.

[24] X. Zhu, Y. Ju, J. Chen, D. Liu, and H. Liu, "Nonenzymatic wearable sensor for electrochemical analysis of perspiration glucose," *ACS Sensors*, vol. 3, no. 6, pp. 1135–1141, 2018.

[25] G. Yuan, Z. Wang, F. Meng, Q. Yan, and S. Xia, "An overview of human activity recognition based on smartphone," *Sensor Review*, vol. 39, no. 2, pp. 288–306, 2019.

[26] M. Szewczyk and M. Kuniewski, "Controlled voltage breakdown in disconnector contact system for VFTO mitigation in gas-insulated switchgear (GIS)," *IEEE Transactions on Power Delivery*, vol. 32, no. 5, pp. 2360–2366, 2017.

[27] C. Du, D. Zou, T. Liu, and W. Li, "A study on the influence of stage load on health monitoring of axial concrete members using piezoelectric based smart aggregate," *Applied Sciences*, vol. 8, no. 3, 2018.

[28] Q. Meng, Y. Chen, M. Zhang, L. Han, H. Pu, and J. Liu, "On the kaiser effect of rock under cyclic loading and unloading conditions: Insights from acoustic emission monitoring," *Energies*, vol. 12, no. 17, p. 3255, 2019.