Real Time Embedded Programming

ENG5220

Course REPORT

2021

Gesture Recognition Using Raspberry Pi

Abstract

The gesture recognition system designed in this paper takes STM32 single chip microcomputer and raspberry pi as the core control unit, 8 copper clad plates as sensors to obtain the capacitance corresponding to the gesture, and uses TI company's sensor chip FDC2214 to detect the capacitance, and uses machine learning KNN (proximity) algorithm for data training and testing to finally recognize the specified gesture. Through the key module and OLED display to achieve human-computer interaction function, it can realize the function of key parameter adjustment and mode switching. The model of server and client based on TCP protocol enables users of raspberry pi to communicate with each other. The whole system involves gesture recognition test structure design, system circuit design, data transmission, data storage and KNN algorithm. Through a large number of tests, the advantages and disadvantages of different schemes are compared, and the existing appearance test structure and control system structure are finally determined. The experimental results show that the design of gesture recognition system is feasible.

Key words: FDC2214, machine learning KNN, microcomputer, TCP protocol, attitude calculation

Contents

1 Introduction

This project uses capacitor boards as a sensor to recognize the voltage and current changes when users make different gestures, and then transmits the data to raspberry pi, which is used for gesture recognition.

When we can achieve this, it will be able to do many things. For example, two people can play with Rock-Paper-Scissors remotely and enjoy an immersive experience through raspberry pi.

As shown in the Figure 1, when player 1 makes a gesture in the area made of capacity boards, the sensor can sense the corresponding voltage and current, and transmit the signal to raspberry pi, and then raspberry pi will be able to recognize the gesture through the acquired signal. At the same time, player 2's gesture signal can be recognized by raspberry pi 2 and transmitted to raspberry pi 1. At this time, raspberry pi only needs to judge the outcome through two signals and output the result on the screen, and a round of the game is completed.

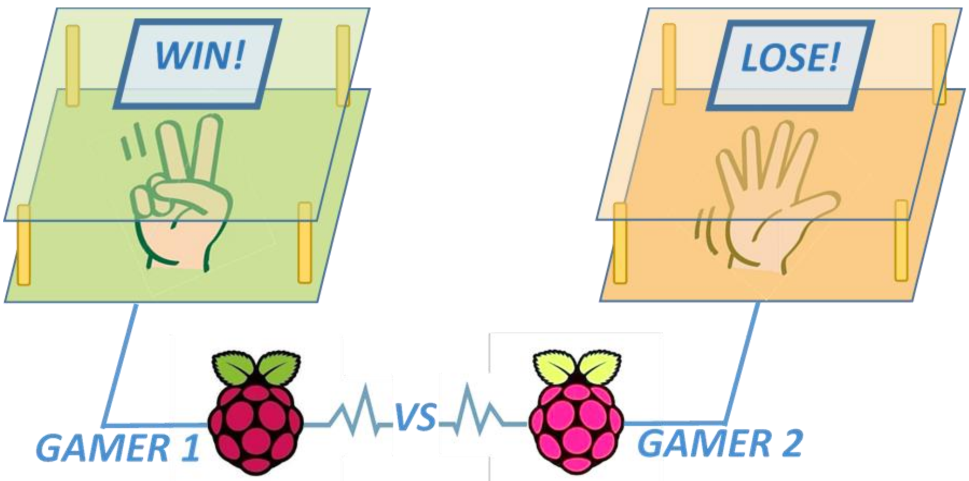

Figure 1 Rock-Paper-Scissors game

In order to achieve this goal, we plan to obtain gestures from different people and gestures' relevant meanings [1]. We will upload several capacitive boards on acrylic plate as well as the Single Chip Micyoco(SCM) to get the gesture signal. Then the figures will be transform to Raspberry Pi to train the machine to distinguish most of gestures by machine learning or deep learning model. We will choose the better model to use. After that, various gestures are able to be recognized by our Raspberry Pi immediately [2]. Finally the two devices exchange collected data and make a decision to be shown on the screen.

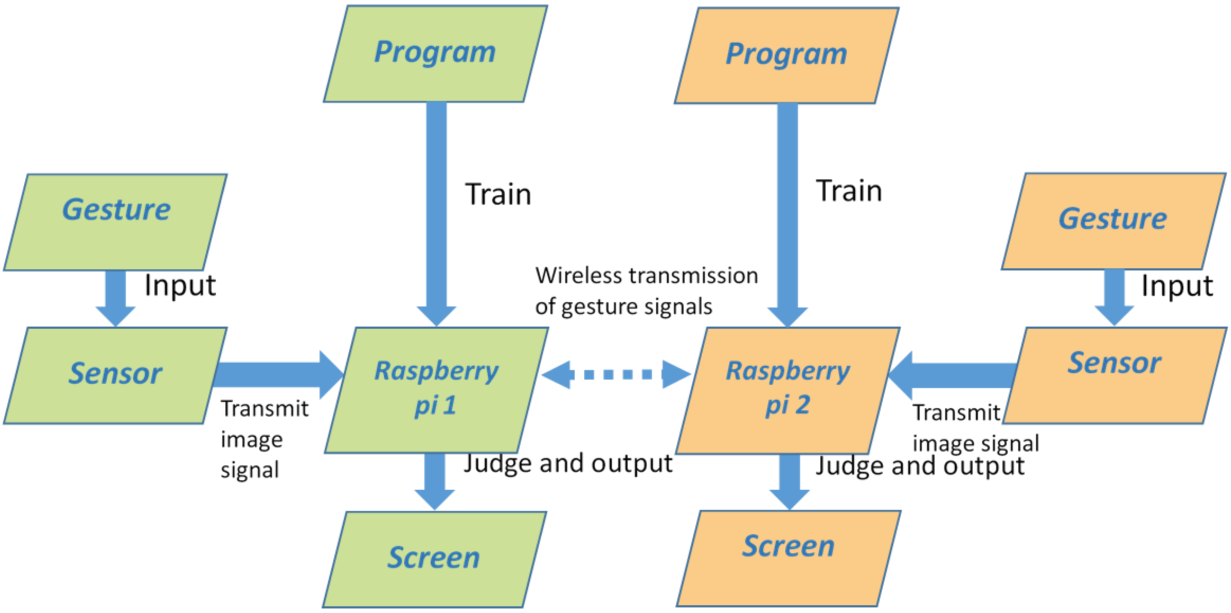

Figure 2 Project design drawing

Of course, the significance of this technology is far more than just playing rock-paper-scissors game [3]. For example, in dark environment where quietness is needed， such as in anti-terrorist CQB operations， it helps to better recognize gesture commands. Besides, in the live stream of virtual youtuber, it is sometimes necessary to capture facial expressions and gestures, but the cost of motion capture equipment is very expensive. In this case, our products will have a good play. The cost of our gesture recognition is very low, so it can support low-cost live stream very well.

2 Aims and objectives

The aim of this project is to achieve remote interaction of gesture recognition. When two people use two raspberry pi to make rock, scissors or paper gestures, raspberry pi can obtain signals from both sides through capacitance sensor and wireless network, identify gestures and judge whether they win or lose.

To achieve the aim above, the objectives are:

1. Using FDC2214 chip to build sensing area to collect data.
2. Design a program, so that the SCM can receive the signal from the sensor and transmit the signal to raspberry pi.
3. Using KNN algorithm to design a program, so that raspberry pi can determine what gesture the received signal represent.
4. The server and client model of TCP protocol are adopted to make the raspberry users communicate with each other.

3 Theory

This system is mainly composed of micro controller, power supply module and capacitance sensor module. The following demonstrates the selection of system and module.

3.1 System structure

In the setup of this system, the PCB with the size of 20cm * 20cm is the measurement area composed of FDC2214 chip and four capacitance plates. FDC2214 is connected to SCM by wire. There is an OLED display on the SCM. Raspberry pi connects to SCM and powers it. The specific appearance test structure is shown in Figure 3.

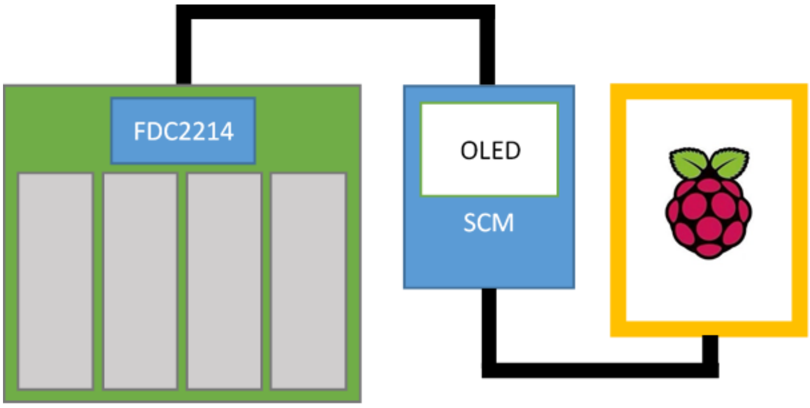

Figure 3 System mechanical structure

3.2 Control scheme and demonstration

3.2.1 Selection and demonstration of micro controller

In this system, a microcontroller is needed to control the whole system, read the data of FDC sensor, and calculate the algorithm, so we need to choose the microcontroller.

Scheme 1: STM32 single chip microcomputer is adopted. 32 single chip microcomputer is cheap and widely used, but its function is single [4]. Even if it can be equipped with SD card to expand the memory, the running memory is only 48KB [5]. As the use of KNN algorithm and even deep learning data, the amount of calculation is too large, and the memory is not enough to use.

Scheme 2: The scheme of using raspberry pi of microcomputer and STM32 single chip microcomputer as controller [6]. The memory of raspberry pi is much larger than 32 single chip microcomputer, which can process complex data and has good computer programming ability. With 32 MCU control switch, OLED and simple data processing can meet our needs [7].

According to the requirements of the project, we choose the second scheme.

3.2.2 Selection and demonstration of capacitance sensor module

In this system, we need to accurately measure the gesture, so we use the capacitance sensor, through FDC2214EVM can accurately collect the capacitance between poles, as shown in Figure 4, each FDC2214EVM has four channels CH0-CH3, and each channel has two interfaces, one of which is ground. We need to evaluate the number of capacitive sensors [8].

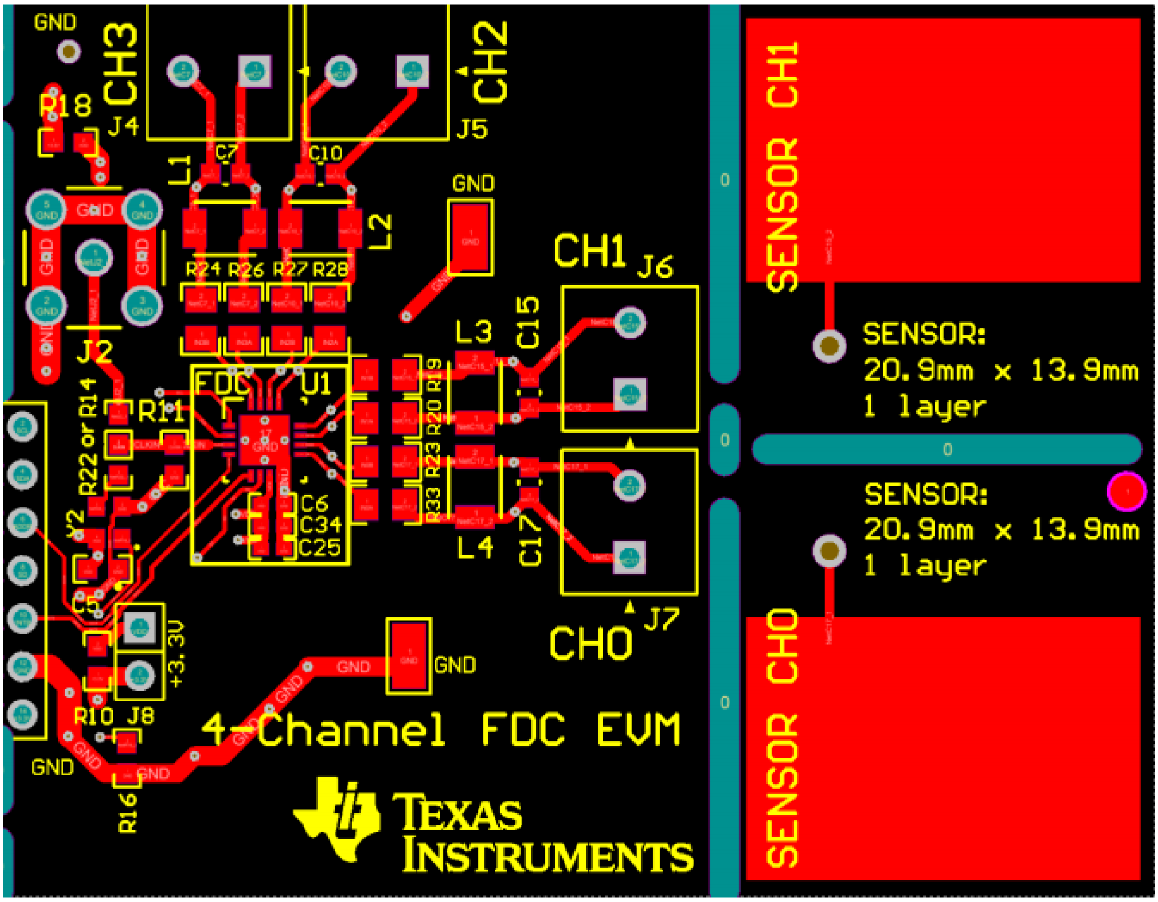

Figure 4 Circuit diagram of FDC2214EVM [9]

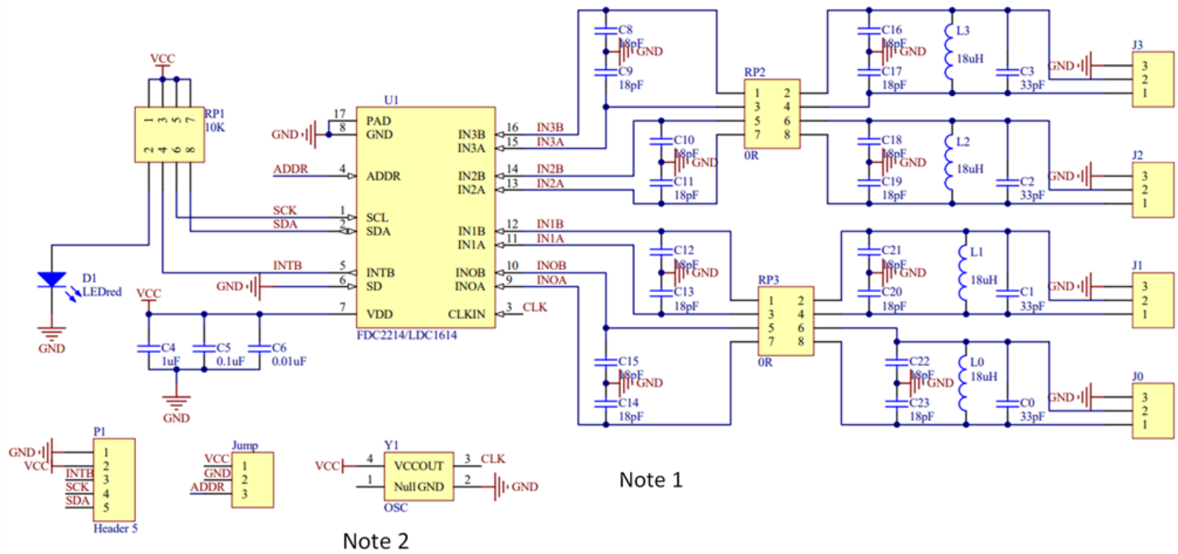

Figure 5 FDC2214 wiring diagram

Note 1: By default, the 18pF capacitor on the reverse side of PCB is not soldered. According to the device manual, the function of 18pF capacitor is to remove the influence of long line effect of transmission wire.

Note 2: Internal clock is used by default, Y1 (40m crystal oscillator) is not welded. CLK is grounded by solder jumper.

As shown in Figure 6, using four capacitor modules, it needs to occupy four channels [9]. Only one FDC2214EVM is needed, with low cost, simple connection and less data to be processed. After testing, STM32 single chip microcomputer can be used to process. It can accurately recognize the gestures not limited to rock-paper-scissor gestures and digital gestures, and will not cause any problems because of the small amount of data.

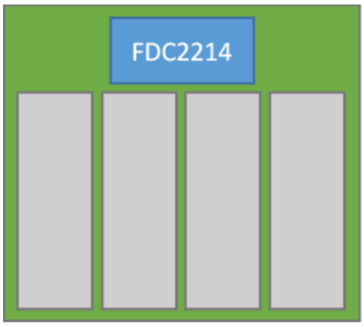

Figure 6 Sensor module

3.3 Selection and demonstration of algorithm

In this system, we need to design an algorithm to process the data of the FDC sensor, choose the appropriate algorithm according to the better results, and finally choose two schemes from the problem design and experience.

Scheme 1: Machine learning KNN algorithm [10]. As a nonparametric classification algorithm, KNN algorithm is very effective and easy to implement [11]. By finding the k nearest neighbors of a sample and assigning the average value of the attributes of these neighbors to the sample, the attributes of the sample can be obtained.

Scheme 2: Using deep learning neural network algorithm. After repeated training, it can achieve a more accurate degree than KNN algorithm. Through retraining, it can strengthen the learning of the established model and identify the users accurately.

According to the requirements of this project, we choose the first scheme

4 Design

4.1 Circuit design

4.1.1 General block diagram of system circuit

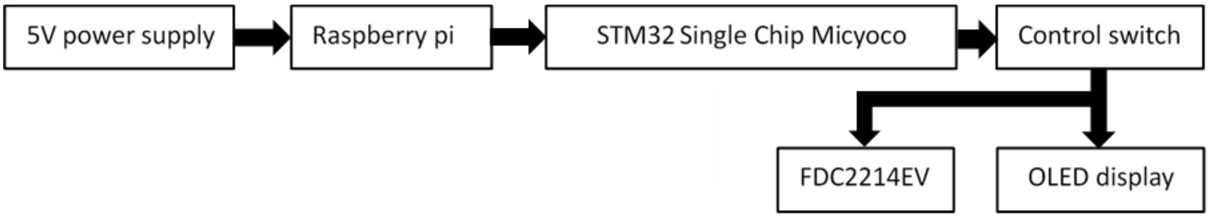

Figure 7 System circuit

As shown in Figure 7, the raspberry pi is powered by a 5V power supply, STM32 is connected to the raspberry pi, and the sensor FDC2214 and OLED display screen are connected to the SCM through the control switch, so that the sensor can transmit the signal to the raspberry pi through the SCM.

4.1.2 Design of transfer circuit

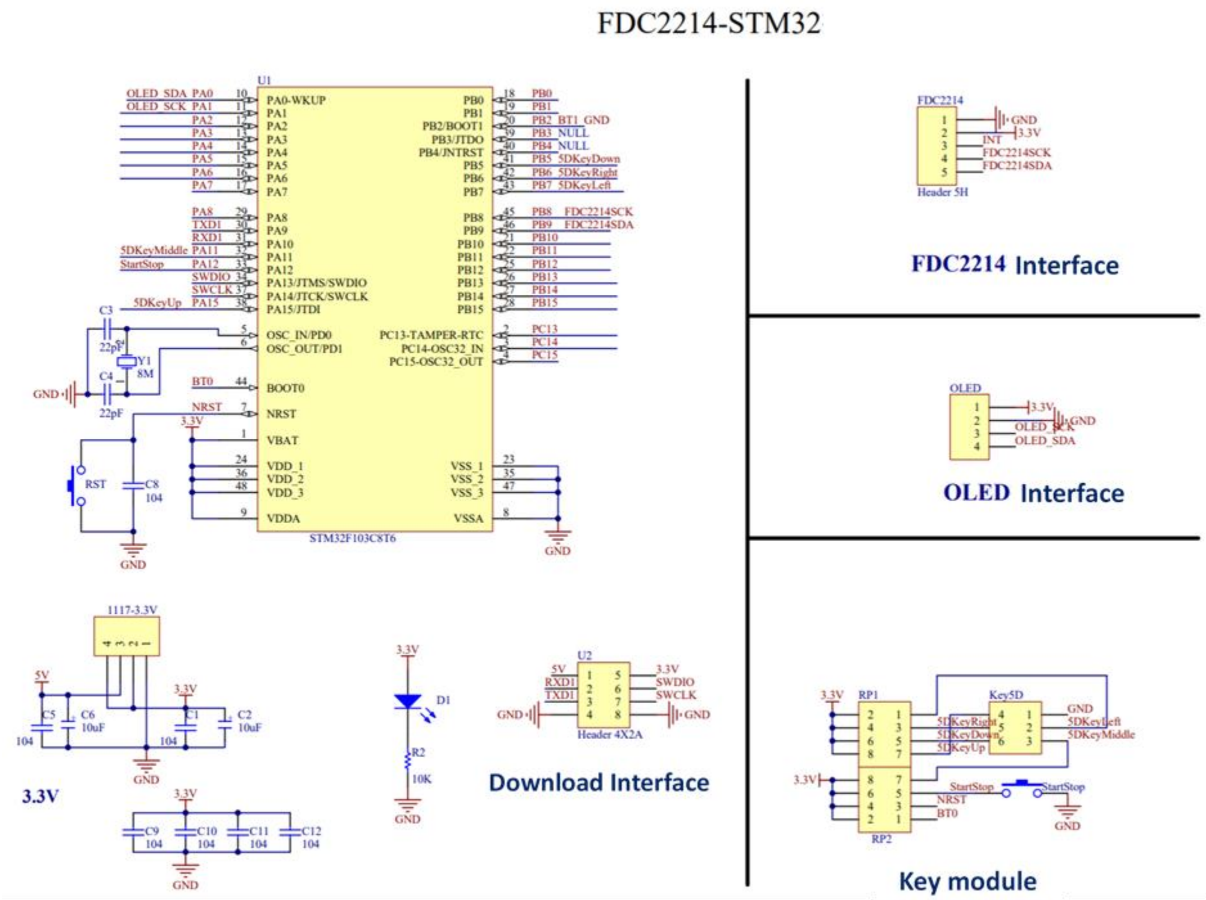

Figure 8 Schematic diagram of transfer circuit

4.2 Program design

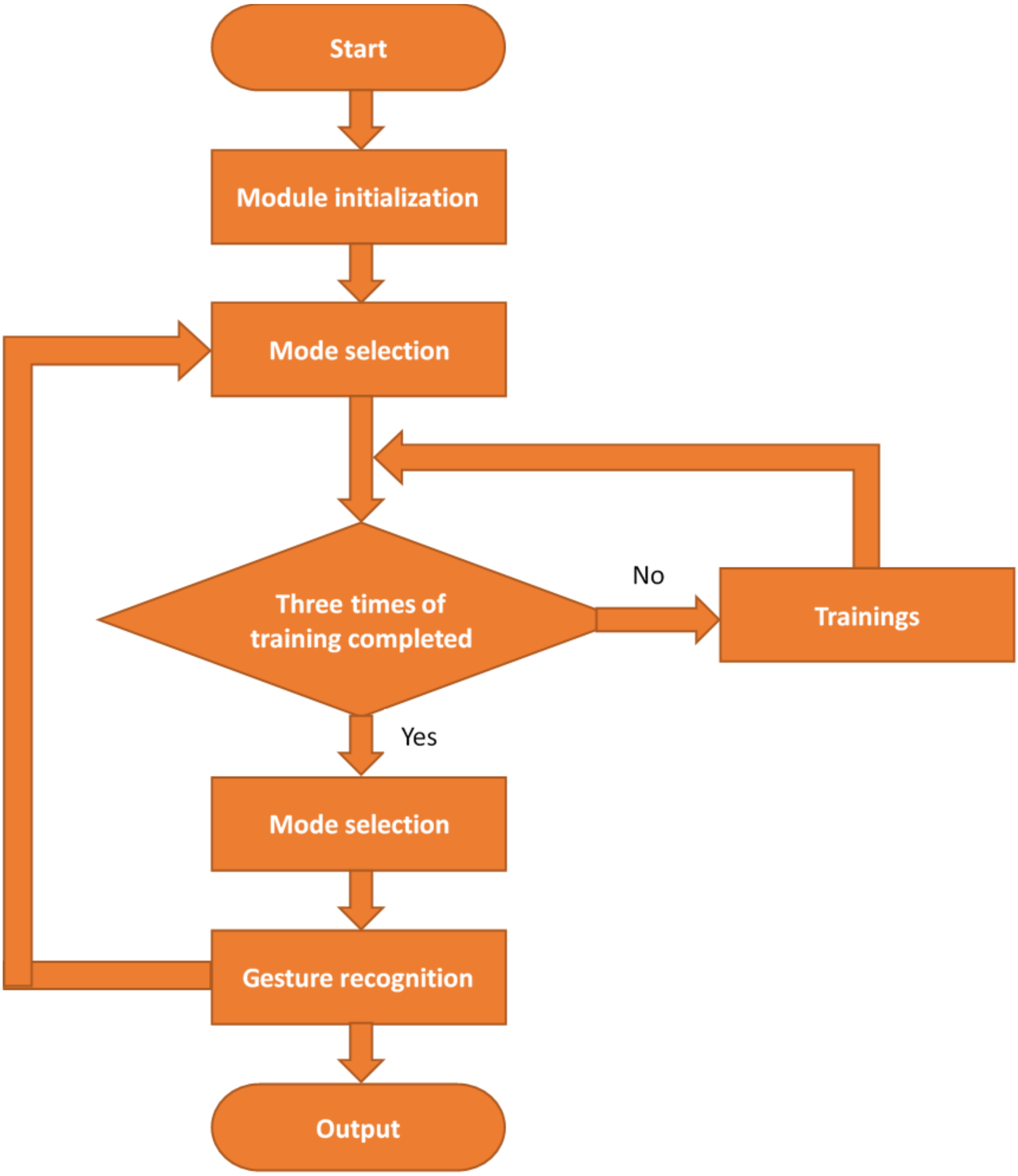

Figure 9 Software flow chart

This program is divided into two parts, one part is to use STM32F103C8T6 MCU to collect FDC2214 sensor data and human-computer interaction part, the other part is to use raspberry pi as data analysis and processing to get the expected results.

STM32 program includes two parts: the first part is initialization peripheral, including OLED initialization, FDC2214 initialization, key initialization, serial communication initialization; the second part is human-computer interaction part, including selection training or recognition, which is roughly divided into five parts There are three categories, which are guessing training part, guessing test part, rowing training part, rowing test part and clearing data set part. The raspberry pi program mainly deals with the signal from STM32 serial port for data receiving training or recognition.

4.2.1 FDC2214 configuration

Table 1 FDC2214 configuration

Register	Value	Register	Value	Register	Value	Register	Value
0x08	0x04D6	0x09	0x04D6	0x0A	0x04D6	0x0B	0x04D6
0x0C	0x1000	0x0D	0x1000	0x0E	0x1000	0x0F	0x1000
0x10	0x000A	0x11	0x000A	0x12	0x000A	0x13	0x000A
0x14	0x1001	0x15	0x1001	0x16	0x1001	0x17	0x1001
0x1E	0x7800	0x1F	0x7800	0x20	0x7800	0x21	0x7800
0x19	0x0000	0x1B	0xC20D	0x1A	0x1E81		

4.2.2 KNN algorithm

Distance calculation method: Euclidean distance calculation method

Table 2 Value of k

K value in rock-paper-scissor recognition mode	K value in digital recognition mode
15	41

1. Calculate distance: Given the test object, calculate the distance between it and each object in the training set.
2. Neighbor finding: The nearest K training objects are delineated as the nearest neighbors of the test objects.
3. Classification: According to the k nearest neighbor belonging to the main category, to test object classification

4.3 Communication module

Communication between devices is required. In our design, the device will communicate in the same LAN. We use the server and client model of TCP protocol [12]. Considering that there are multiple clients communicating with the server at the same time, we use the fork() function to create multi process TCP connections. Because the data uploaded by the clients connected by each process needs communication comparison, we use pipe () to communicate among the processes. Taking the rock-paper-scissors game as an example, each client sends the result of the user gesture recognition to the server. Each process of the server is responsible for receiving data from a client. In the judgment section, the process needs to know how much data the clients responsible for other processes send, so they use pipe to transfer data to each other. After the judgment is completed, each process returns the judgment results to the client in charge. The structure of the whole communication module is shown in Figure 10.

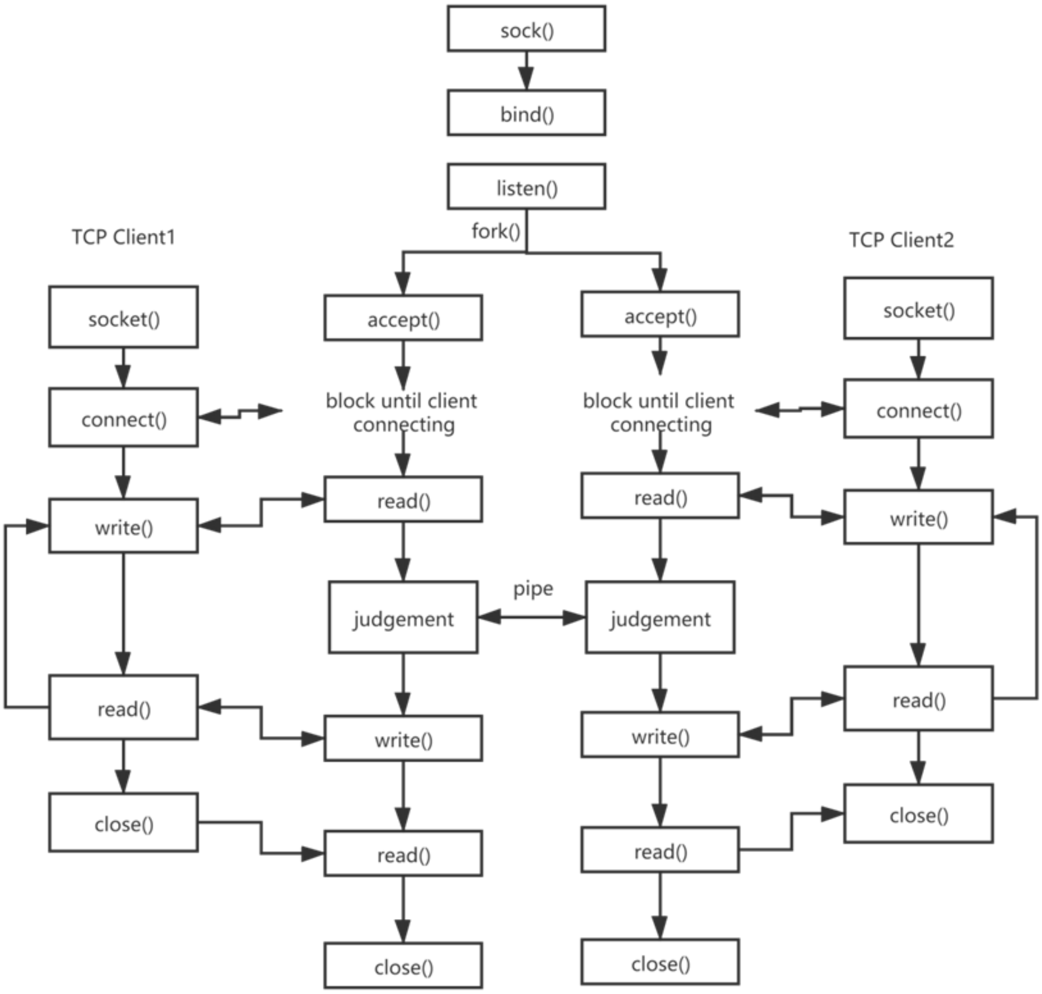

Figure 10 The structure of communication module

5 Results and critical discussion

5.1 Gesture recognition

5.1.1 Test plan

The test plan is black box test, no screening, looking for any tester, according to the order of routine test, pressure test and acceptance test. Each round of the test is divided into 24 training sessions, including 9 rock-paper-scissors (stone, paper, scissors) training sessions and 15 digit recognition (0, 1, 2, 3, 4, 5) training sessions. After each round, the candidates are replaced.

5.1.2 Test result

1. Routine test:
   Test plan: Place the hand in the central normal area, and move it slightly during training and test.

   Table 3 Routine test

Number of test rounds	1	2	3	5	10
Wrong number of mode 1	0	0	1	2	4
Wrong number of mode 2	1	3	4	5	9

   Conclusion: The test can meet the requirements of the topic, there is a slight error, and the accuracy rate is 95%.
   Note: There are 24 experiments per round.
   Mode 1: Rock-paper-scissors recognition
   Mode 2: Digit (0, 1, 2, 3, 4, 5) recognition

2. Stress test:
   Test plan: Place the hand in the sensing area and try to move to the edge of the sensing area during training and testing.

   Table 4 Stress test

Number of test rounds	1	2	3	5	10
Wrong number of mode 1	2	3	3	6	12
Wrong number of mode 2	4	10	16	28	58

   Conclusion: There are some errors, basically because the number of training times is too small, the error area is almost the untrained area of the sensing edge, and the accuracy rate is only 70%.
   Note: There are 24 experiments per round.
   Mode 1: Rock-paper-scissors recognition

Mode 2: Digit (0, 1, 2, 3, 4, 5) recognition

3. Acceptance test:

   Test plan: Select any person who has not been tested at all, place the hand in the test area close to the central position for training, and move slightly during the test but try not to touch the sensing edge area.

   Table 5 Acceptance test

Personnel number	1	2	3	4	5	6
Test times of mode 1	9	9	9	9	9	9
Wrong number of mode 1	1	0	1	1	0	0
Test times of mode 2	20	20	20	20	20	20
Wrong number of mode 2	4	5	3	4	5	4

   Conclusion: Through the test can meet the requirements of the topic, but in the training did not touch the place, the accuracy still needs to be improved, and the accuracy rate is about 85%.

   Note: Mode 1: Rock-paper-scissors recognition

   Mode 2: Digit (0, 1, 2, 3, 4, 5) recognition

5.1.3 Test analysis

On the whole, the system achieves good performance. After a limited number of training, it can recognize the corresponding gesture, and the error is small. However, the judgment error is relatively large when the number of training is less, and the position of the gesture has a greater impact on the judgment, so the structure still needs to be optimized.

5.2 Communication

5.2.1 Result of communication part

Test result of communication part is as follow:

Figure 11 is the running result of the server. Through this screen capture, we can see what data is sent by each connected client, and judge the result.

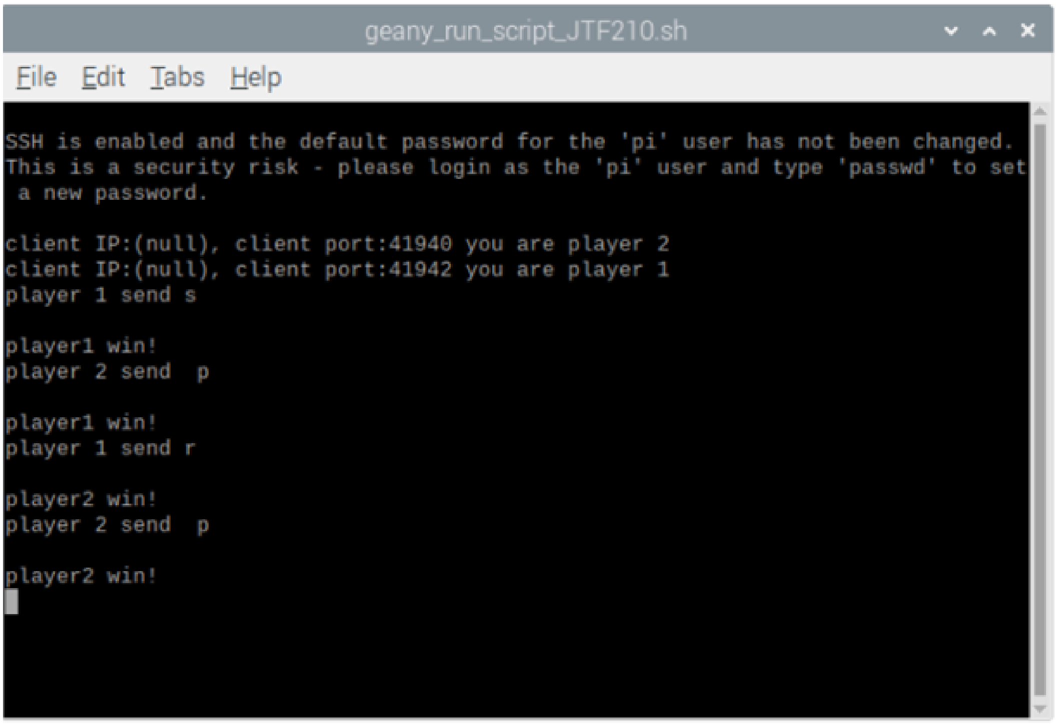

Figure 11 Running result of the server

Figure 12 and Figure 13 are the running results of two clients. Where p is paper, R is rock and S is scissors. We can see the judgment results from the server.

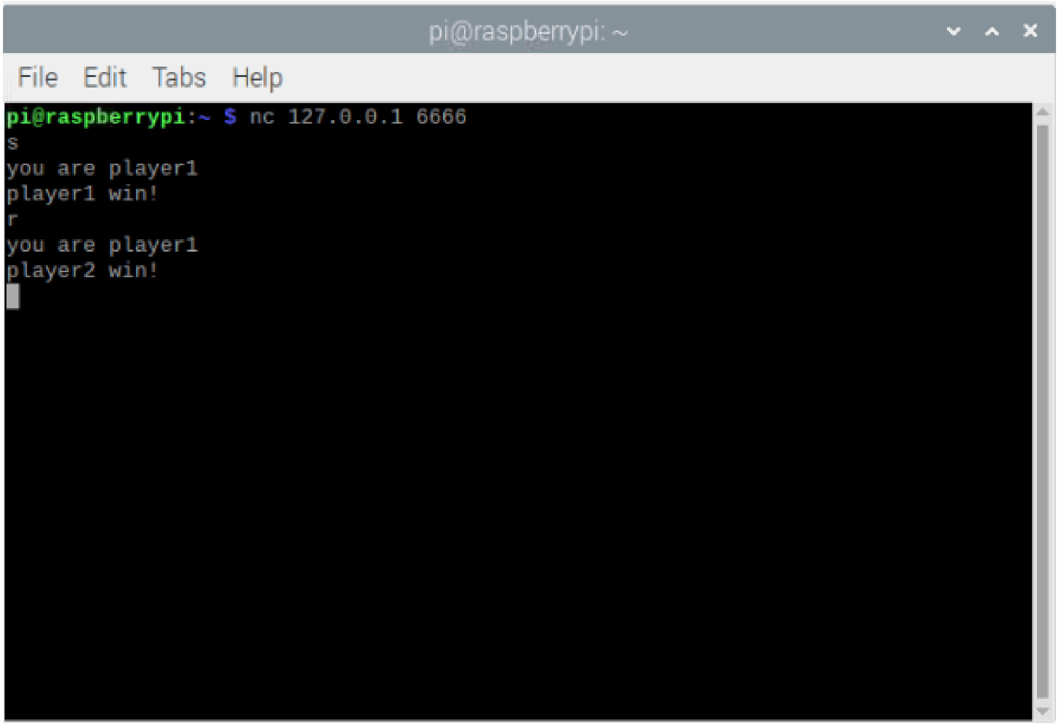

Figure 12 Running results of client 1

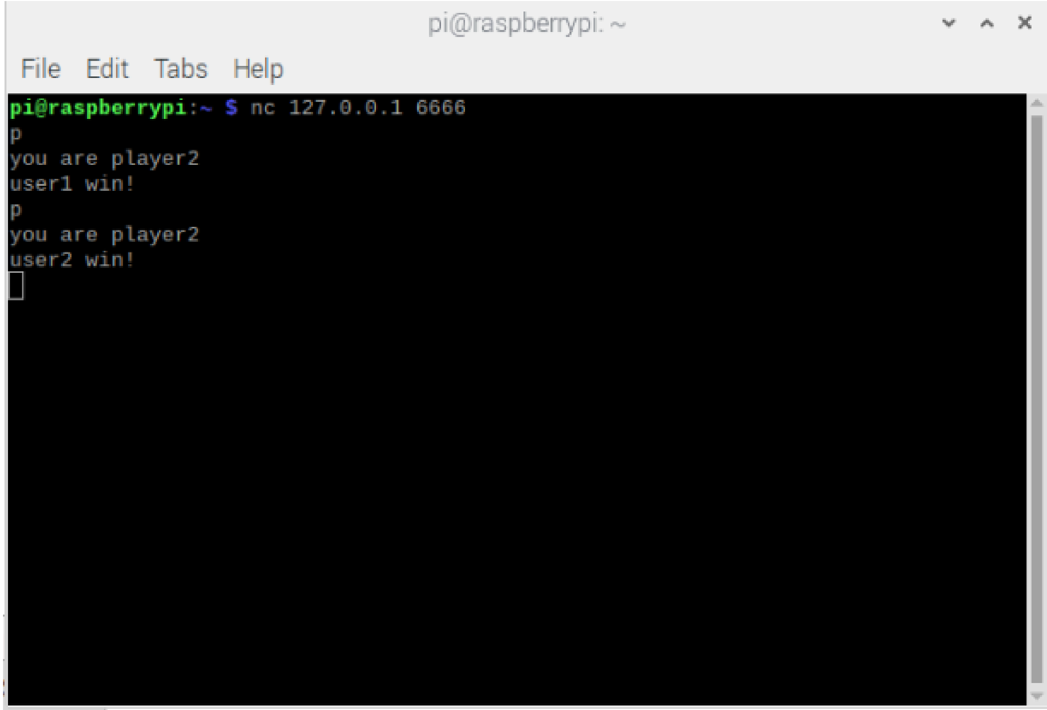

Figure 13 Running results of client 2

5.2.2 Analysis

Achieve the goal successfully. Two players can play rock paper scissors game remotely through two raspberry pies. Raspberry pies will collect signals from both sides at the same time for gesture recognition and judgment, and then feedback the information to players.

But originally we planned to use QT to complete the online part, and the window has been designed, but due to some unknown technical problems, the signal cannot be connected. The replacement scheme causes the online communication only in the LAN, and cannot communicate with Wan. We will continue to study this in the future.

6 Conclusions

In this project, we have completed the expected plan. When two people use two raspberry pi to make rock, scissors or paper gestures, raspberry pi can obtain signals from both sides through capacitance sensor and wireless network. At the same time, we have also completed more content, so that raspberry pie can recognize not only the stone, paper and scissors gestures, but also the number 0 to 5 gestures. The purpose of adding such a function is to realize the situation described in the first chapter in the future, which can be used for encrypting remote calls. Of course, this is only a preliminary result, proving the diversity of gesture recognition using raspberry pi.

In arriving at this result, the following work have been done:

1. Build gesture recognition platform. In this process, we choose to use a PCB with the size of 20cm * 20cm is the measurement area composed of FDC2214 chip and four capacity plates to build the sensing area. At the same time, FDC2214 is connected to SCM by wire. There is an OLED display on the SCM.
2. The program makes STM32 MCU can collect the data of FDC2214 sensor and transfer it to raspberry pi. At the same time, it can complete the human-computer interaction part through the peripheral OLED, including selecting mode, clearing data and so on.
3. KNN algorithm is used to realize gesture recognition using the signal obtained by raspberry pi. Through training, raspberry pi can now use two patterns to recognize gestures, one is to recognize the gestures of rock, paper and scissors in the game of stone-paper-scissors, the other is to recognize 0 to 5 Digital gestures.
4. The model of server and client based on TCP protocol enables users of raspberry pi to communicate with each other. This enables us to complete the initial goal. Two players can play rock-paper-scissors game remotely through two raspberry pi. Raspberry pi will collect signals from both sides at the same time for gesture recognition and judgment, and then feedback the information to the players.

Future work:

At the same time, the project still has some unfinished and worthy of further study, which is worthy of further study in the future. Originally, we planned to use QT to complete the online part, and the window has been designed, but due to some unknown technical problems, the signal cannot be connected, so we can only regret to abandon this scheme temporarily. As a result, we can only communicate in LAN, not wan. We will continue to study this in the future.

At the beginning of the project, we proposed that gesture recognition technology can be used in many scenes, such as motion capture or encrypted communication. To achieve these goals, there are still higher requirements for the recognition ability of raspberry pi. Although we have done more patterns in the project to complete more diverse identification, it is still far from enough. There is still much room for improvement in this aspect in the future.

7 References

[1] Rautaray S, Agrawal A, Vision based hand gesture recognition for human computer interaction: a survey, ARTIFICIAL INTELLIGENCE REVIEW, 2015, 43: 1.

[2] Bretzner L, Laptev I, Lindeberg T, Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering. Fifth IEEE intern. Conf. Auto. Face Gesture recognition, 2002: 405.

[3] Kevin N, Ranganath S, Ghosh D, Trajectory modeling in gesture recognition using cybergloves and magnetic trackers, IEEE TENCON, 2004, 10:571.

[4] STMicroelectronics, Embedded Pi, triple-play platform for STM32, Raspberry Pi and Arduino, 2016.

[5] Knorig R, STM32 Development with CubeMX and Eclipse, Raspberry Pi Geek, 2017, 22.

[6] STMicroelectronics, RM008 Reference Manual：STM32F101xx, STM32F102xx, STM32F103xx, STM32F105xx and STM32F107xx advanced Arm-based 32-bit MCUs, 2021.

[7] Yiu J, System-on-Chip Design with Arm Cortex-M Processors, ARM Education Media, 2019.

[8] Texas Instruments, Capacitive Proximity Sensing Using FDC2x1y, SNOA940A, 2017.

[9] Texas Instruments, FDC2x1x EMI-Resistant 28-Bit, 12-Bit Capacitance-to-Digital Converter for Proximity and Level Sensing Applications, SNOSCZ5A, 2015.

[10] M. Parsian, Data Algorithms, O'Reilly Media, Inc. 2015.

[11] M. Chatterjee, A Quick Introduction to KNN Algorithm, 2020.

[12] C. Hunt, TCP/IP Network Administration, O'Reilly Media, Inc. 2002.

[13] Cote M, Payeur P, Comeau G, Comparative study of adaptive segmentation techniques for gesture analysis in unconstrained environments, IEEE inter. Workshop Imagining Systems Techniques, 2006: 28.