A Cyber Monitoring System Based on Image Recognition

Shi-Xun Huang
Department of Electronic Engineering,
National Chin-Yi University of Technology
Taichung, Taiwan
pu758614@gmail.com

Wen-Yuan Chen
Department of Electronic Engineering,
National Chin-Yi University of Technology
Taichung, Taiwan
cwy@ncut.edu.tw

***Abstract*—This study is aimed at the network monitoring system for home use and uses video recognition methods to determine whether the surveillance area is intruded or not. The user can receive notifications by using the instant push function.**

The hardware is developed using Raspberry pi 3, and OPENCV is used as the software in the image office to identify objects or persons with motion detection and face recognition. When a person or an object passes through, the push function is activated, and the message is notified to the user by using LINE and e-mail; and the web interface is provided so that the user can instantly monitor and control the direction of the shot, and view all event information in the whole process.

Compared to using motion images on the market for network monitors to continuously record and rewrite memory cards, this study using image recognition to record real-time images can significantly reduce memory usage and increase the duration of data storage. And through the popular LINE instant messaging APP, e-mail, and web interface, users can get started faster without having to familiarize themselves with new applications.

Keywords-Internet of Things; Raspberry Pi; Image Detection; Cloud

I. INTRODUCTION

The Internet of Things (IoT) [1], an important technology for Industry 4.0, was mentioned in "Future of the Future" [2] written by Bill Gates in 1995. At the World Summit on the Information Society(WSIS), the "ITU Internet Report 2005: The Internet of Things"[3] released by the International Telecommunication Union put forward the specific concept of the Internet of Things and officially opened the era of the Internet of Things.

The Internet of Things [4] has been applied in the areas including homes, factories, cities, hospitals, etc.. In 2011, the German government proposed "Industry 4.0" and officially opened a new generation of industrial revolution. The Internet of Things has changed the human's life styles, and makes us live more conveniently and efficiently.

Everyone values the safety of their own personal belongings. When the family was robbed by a thief or had a car accident that is escaped, the first reaction was to visit the intersection monitor or obtain a driving recorder for nearby vehicles so to find the truth of the field and clarify the responsibility for the accident. Therefore, the "image record" has obviously become one of the most favorable evidences in the incident.

With the advancement of science and technology and the need for human security, the number of monitors has exploded. From closed-circuit surveillance cameras on traditional roads to cameras in driving vehicles, the recorder has become one of the necessary equipment. In recent years, monitors for home use have become more and more popular, and they all use network cameras [5] because they are easy to install and can be networked with real-time monitoring functions such as real-time monitoring and theft prevention. Therefore, network monitors for home use have gradually become the necessary home equipment, so it can be seen that network monitors for home use have become the mainstream devices for people to pursue home security.

II. SYSTEM STRUCTURE

A. *Organization and Hardware Architecture*

This system uses Raspberry pi 3 as shown in Fig. 1 for research and development. Raspberry pi 3 is an embedded development board based on Linux operating system. It is a multiplexed processor with many communication interfaces that can meet this requirement for research development needed.

Fig. 1: Raspberry pi 3

In order to allow users to control the image angle, in this study, a mechanism was designed using AutoCAD as shown in Fig. 2, and the entity was printed using 3D printing. This mechanism can load servo motors and USB webcams.

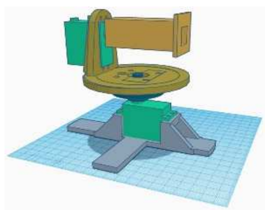


Fig. 2: Servo Motor and USB Network Camera Loading

The servo motor uses MG996. There is a variable resistance sensor inside the motor to sense the rotation angle. As long as pulse width modulation is used to release the specified value, it can be rotated to the desired angle according to the width of the PWM signal high potential, which determines the servo motor rotation angle, as shown in Fig. 3.

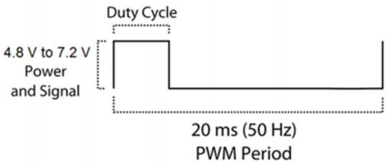


Fig.3 : MG996 signal control

Because Raspberry pi 3 voltage is only 3.3V, it is not enough to drive the servo motor; therefore, PCA9685 module which can raise the voltage to 5V was used, and I2C is used to connect the module and control PWM output. The hardware architecture diagram is shown in Fig. 4, and the actual hardware diagram is shown in Fig. 5.

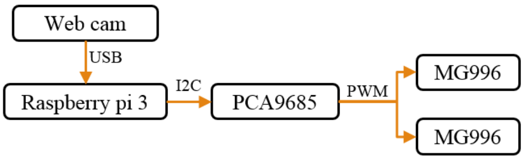


Fig.4 : Hardware architecture

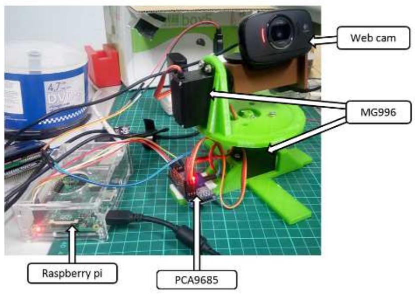


Fig.5: Organization hardware entity map

B. *Image Recognition*

The image recognition part uses two methods: motion detection and face recognition. Because only one method is used to monitor, it is easy to have loopholes. In some conditions, it can avoid detection, so two methods are used at the same time to improve the reliability of security.

1) *Motion Detection*

a. *Background subtraction*

By using the image background subtraction method [6] to set the image in the stream, two frames are intercepted. The difference between the two frames captured in this study is 0.3 second. Taking Fig. 6 as an example, the door-to-door image of a person opening the door is obviously changed very much. Therefore, after the subtraction of the image, Fig. 6(c) can be obtained; in contrast, if the door is always static, the environment is not changed much, as shown in Fig. 7(c), and the screen is almost black.

(a) (b) (c)

Fig.6: One of the background subtraction experiments (a) The first frame, (b) The second frame, (c) After the subtraction of the image

(a) (b) (c)

Fig. 7: Background subtraction experiment diagram (II); (a) First frame, (b) Second frame, (c) After image subtraction

b. *Color transformed into grayscale*

The general photo is composed of three primary colors of RGB (red, green, blue). In image processing, color images are often converted to grayscale (black and white) to increase the efficiency of calculation. We then apply this method to the images computed in the previous section. The result is shown in Fig. 8. The color (a) diagram is converted to the grayscale (b) diagram.

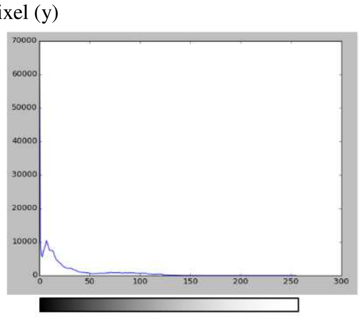
(a) **(b)**

Fig. 8. RGB to grayscale; (a) RGB image (b) Grayscale image

c. *Grayscale image histogram*

Fig. 8 (b) is drawn as a histogram, and the result is shown in Fig. 9 below. The x-axis is the gray level, and the y-axis is the total pixel. Obviously, the more black (0) the total pixels, the more negative 255 (The smaller the total number of white pixels, the larger the black portion and the smaller the white portion).

Pixel (y)



Grayscale (x)

Fig.9: Grayscale histogram of Fig. 8(b)

d. *Result analysis*

Next, the above experimental method was conducted at the same location but at different time points. The results are shown in Fig. 10.

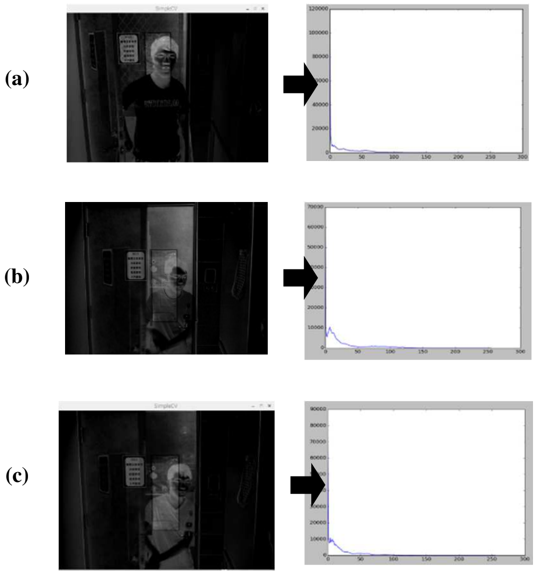
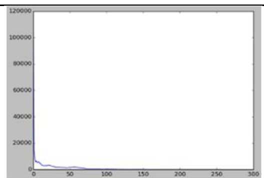
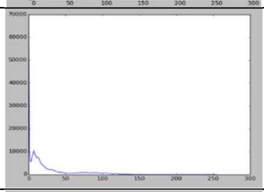
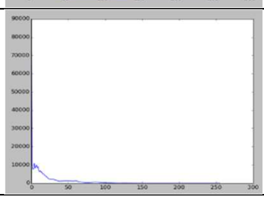


Fig.10: Histogram analysis

Analyzing the values of the histogram, we find that the gray level is stable at about 26 to 75, so the total pixel percentage of 25 to 75 grays is calculated. As Formula (1), the calculated percentage can also be regarded as the change value of the screen. The higher the change in the picture, the greater the amplitude, as shown in Table 1.

$$n_{26}\sim n_{75}\ \text{percentage} = \frac{n_{26}+n_{27}+n_{28}+\cdots n_{75}}{n_1+n_0+n_3+\cdots n_{255}} \times 100\% \qquad (1)$$

TABLE I: Histogram Grayscale 25 to 75 Total Pixels Percentage

Histogram	**Percent of 25 to 75 grays**
	26.51%
	17.51%
	19.33%

2) Face Recognition

In this study, face recognition uses OPENCV's built-in facial recognition function library. OPENCV's face recognition method uses Haar's feature method [7]. In the library, the trained cascade is used. The classifier, which only needs to bring the test image into the function, is used to filter the feature. Once this feature passes the screening of all classifiers, the area is determined to be a human face.

In the flow as shown in Fig. 11, the image is subjected to Haar feature facial recognition. When the image passes through the n-type feature classifier, it is determined as a human face.

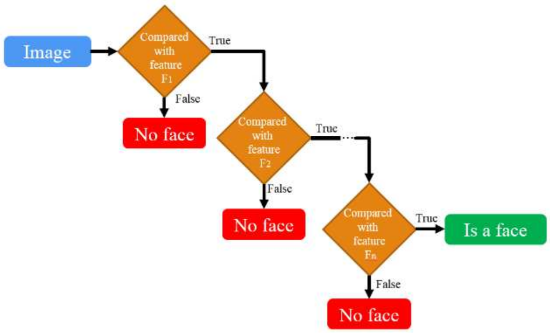


Fig.11: Hart identification method flow chart

C. *Instant Messaging*

The cloud monitor is an indispensable push reminder. Cloud monitors in the general market usually provide their own set of APPs. However, this research is based on the LINE application, which is used by over 90% of current Taiwanese users and is a universal e-mail approach.

a) LINE

LINE has an official account number "Line Notify", which provides APIs for developers to use, and can be developed in Python, C++, C#, and PHP and other programming languages, so it is easy to combine other application features. Users can add the Line Notify account to their own group, and then use the program to connect to Line Notify's API to control sending messages, pictures and official maps.

The transmission process is as shown in Fig.12. The image cannot be directly transmitted from the host, and it must be placed in a public space on the Internet. Therefore, a picture cloud space called "imgur" is used. You can use the imgur's API to enable users to upload pictures via the program, and a series of JSON-formatted messages will be sent back. The message contains the public URL of the picture, so we just send the URL to Line Notify, and we will get the picture through this URL; finally, we will send the message, and the picture is passed to the user’s LINE group.

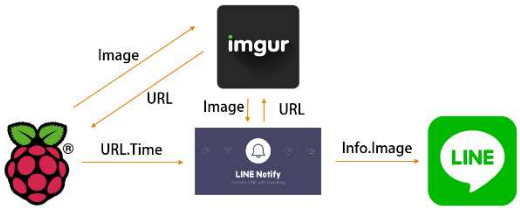


Fig.12: LINE transfer process

b) E-mail

Although LINE is popular in Taiwan and several countries, LINE messages are less likely to be retained for a long period of time. Therefore, e-mails are used to allow users to receive and retain an additional message.

The e-mail in this study was sent via the SMTP (Simple Mail Transfer Protocol) server. When an event was received, the SMTP server was started to send the message packet to the user's E-mail as shown in Fig. 13.

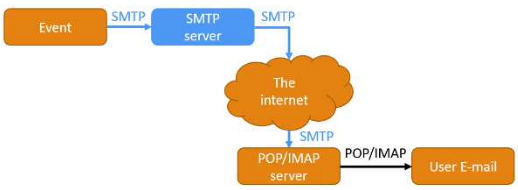


Fig. 13: SMTP Email Flowchart

D. System Integration

In order to integrate all functions and provide a user with a page that can be operated, a webpage platform is designed here. One of the most important features of the web server is that platform support is available in most operating systems.

After entering the page for user information security, the user will be required to enter the user account and password as shown in Fig. 14. After the input is correct, it will directly go to the welcome page as shown in Fig. 15 to represent successful entry into the system.

Fig. 14: Login page

Fig. 15: Welcome page

As shown in Fig. 16, when you enter the Streaming page, you can see the live stream monitoring screen. Users can also choose whether to start Motion detection and Face recognition on the page. On this page, you can also control the rotation of the servo motor through the pull-stem function, and then adjust the lens to your desired angle as shown in Fig.17.

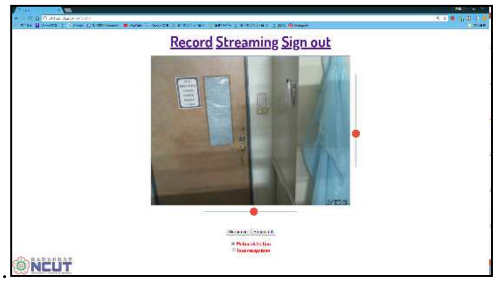


Fig.16: Streaming page

(a) (b)

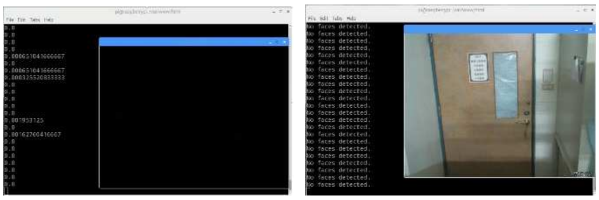
Fig.17: Control the lens angle

The Record screen is the record of historical events. When it is detected that someone or object passes through the monitoring screen, it will not only push to the LINE instant messaging APP and e-mail, but also store the time and image in the database and the host. The user can browse all histories from the web page, as shown in Fig. 18.

Fig.18: Historical event records

III. EXPERIMENTAL RESULTS

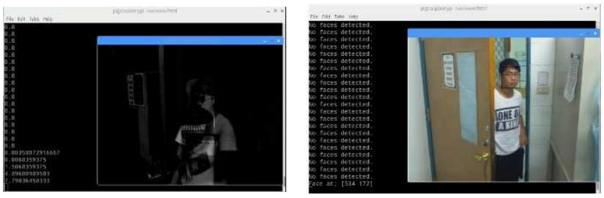
In this experiment environment, the image of the general indoor doorway is monitored, and motion detection and face recognition are enabled at the same time. Therefore, if someone enters or leaves the room, the push system will be detected and started. As shown in Fig. 19, in the case of no movement, the motion detection has almost no change value under the screen, the background subtraction screen is also completely black, and the face recognition is also not detected.



(a) (b)

Fig.19: The picture is motionless; (a) motion detection (b) face recognition

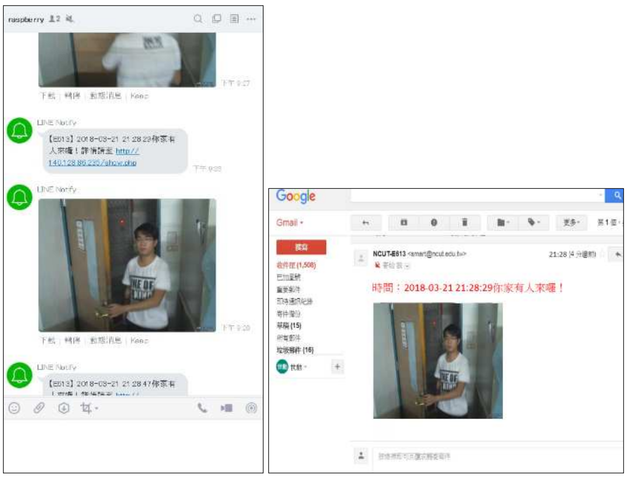
When the door is opened and a person walks in, as shown in Fig. 20, the identified image clearly shows changes in the screen, and the rate of change also increases significantly. In this experiment, if the set value exceeds 5, it is determined that the person is walking in, and the face recognition function also recognizes the face when the door is opened.



(a) (b)

Fig.20: When people come and walk in; (a) Motion detection (b) Face recognition

When the screen change value exceeds 5 or the face recognition recognizes the face, the system immediately activates the push function, and the user can immediately receive the push message from the LINE application and the E-mail as shown in Fig. 21. At the same time, the time information and pictures are stored in the database, and the user can connect to the web server to view as shown in Fig. 18.



(a) (b)

Fig.21: Push broadcasting system; (a) LINE (b) E-mail

IV. Conclusion

Through experiments, the monitor detects the presence of people or objects through the motion detection and face recognition methods. Users can receive notifications via the LINE mobile application and e-mail in real time. In addition, it can be stored in the database in the same manner as the envoys in the form of pictures, which can greatly reduce the more memory usage space and increase the more time for storing data when compared with conventional monitors.

The user can also go to the web server, and just log in the user account to watch the real-time monitoring screen; and the user can adjust the direction of the monitoring angle at the same time, and view the history of all event data page.

References

[1] L.A. Griecoa, A. Rizzoa, S. Coluccia, S. Sicaric, G. Piroa, D. Di Paolab, G. Boggiaa,“IoT-aided robotics applications: Technological implications, target domains and open issues” Computer Communications Vol. 54, pp 32–47,2014..

[2] B. Gates, N. Myhrvold and P. Rinearson, “The road ahead” Viking Penguin Publishing, 1995.

[3] Strategy and P. Unit, “Itu internet reports 2005: The internet of things” International Telecommunication Union (ITU), 2005.

[4] A. Gaur, B. Scotney, G. Parr, S. McClean, “Smart City Architecture and its Applications Based on IoT,” Procedia Computer Science,Vol. 52, pp 1089–1094,2015.

[5] S.N. Jyothi, and K.V. Vardhan, “Design and Implementation of Real Time Security Surveillance System Using IoT,” Proc. of International Conference on Communication and Electronics Systems, Coimbatore, India,October 2016.

[6] M.Piccardi,“Background subtraction techniques:a review,”2004 IEEE International Conference on Systems,Man and Cybernetics,Vol.4，pp。3099-3104，2004.

[7] R. Lienhart and J. Maydt, "An extended set of haar-like features for rapid object detection," in Image Processing. 2002. Proceedings. 2002 International Conference on, pp. I-900-I-903 vol. 1, 2002.