Automatic Video Security System Based on Face-Recognition and Wireless Communication

Ming-Liang Su[1], Zi-Jun Chen[2], Bo-Ya Lai[3], and Zhi-Yong Pang*

School of Physics and Engineering, Sun Yet-Sen University

No. 135, Xingang Xi Road, Guangzhou, 510275, P. R. China

[1]1525491619@qq.com

[2]1534423308@qq.com

[3]273143203@qq.com

*Instructor: 463092989@qq.com

Abstract—**A new kind of automatic video security system based on face-recognition & wireless communication is introduced in this article, including its design method, hardware and software architecture, and its operation results. In this system, the face images captured by the camera will be compared with the criminal face images in the original database, and the similarities between the two faces will be counted. If the similarity is higher than the threshold, an alarm signal will be sent out. The system consists of two parts. The first part is based on FPGA, which includes DE2-70 board offered by Altera corp. and peripherals such as analogical camera & CDMA Modem, etc. First, amount the hardware architecture, modules described by Verilog HDL and SOPC, within which Nios-II processor has been added, are built. Second, amount the software architecture, software is described by C/C++, a μC/OS-II is transplanted & customized tasks have been set. The function of part 1 are face-detection, face image interception, JPEG encoding & MMS (enveloping the JPEG face image) transmission. As to the second part, it is based on the PC program described by C/C++. The only peripheral in this part is another CDMA modem. The function of part 2 are MMS reception, JPEG files extraction & face-recognition (PCA). The system makes use of the advantages of FPGA, such as high operating speed, high integrated level & great convenience, etc. It also makes use of the existing wireless communication network & the PC which is highly efFigient in calculation, compared to FPGA. Therefore, a possible pattern of FPGA—remote wireless communication–PC combined system for video surveillance & intelligent modes-recognition have been introduced.**

Keywords— **FPGA; DE2-70; Nios-II; SOPC, Altera; μC/OS-II; face-recognition; PCA; CDMA; MMS**

I. DESIGN OVERVIEW

A. Purpose

With the development of globalization, most of the countries open their doors to foreign guests. At the same time, security problem is becoming more & more serious under the background of international communication. New technologies are used to improve the performance & efFigiency of the security system. Amount the relevant new technologies, face recognition is a signiFigant one. Video surveillance system using face-recognition technology could be applied to the customs, airports, government securities and other public occasions, assisting in combat against stealing, robbery, terrorism, drug smuggling, human trafFigking and evil religions. That's why we choose to build an ***automatic video security system based on face-recognition and wireless communication*** under the theme of protecting our people, including their lives & wealth.

In practical situations, video surveillance cameras have been installed in every road, every block, every building & every room, etc. We aim at equipping every camera with the ability to recognize human faces and communicate with the police. If the human face, captured by the camera, belongs to a criminal, an alarm signal as well as the location and time of the relevant camera would be send to the local police, assisting them in tracking the criminals. Otherwise, the alert will not be triggered.

However, two practical problems must be solved in this case. First, the database of criminal messages has been installed to the server in the police station, and it is too large for usual ROMs or flash memories to store. In other words, it is unpractical to install huge storages for every camera. Second, millions of faces will expose under the supervising range of one single camera (one face-recognition system) in one day. Therefore, the requirement of each system's computation ability will be extremely high that none of the PC or FPGA system could reach them. Only super server can possibly finish such a tough job.

Since we are not so abundant in hardware resources, we need to find out another way to solve these problems.

B. Features

The solution to these problems, which is also the feature of our design, is that we divide the system into two parts and

assign the first part--video surveillance system connecting a camera--face-detection task while leaving the PCA task to the second part--police station's server. In detail, in the first part, human face images that occur in video will be captured, converted into JPEG format and sent to the terminal—the server of public security bureau, along with the location of the camera & the present time, by using MMS. In the second part, the server will then analyse the received face images & judge whether this face is belong to a criminal. In this way, video surveillance system can become intelligent in face-detection while the server with strong computation ability and the existing wireless communication network can also exert their power, making it possible for computationally expensive real-time face recognition.

C. Applications & Users

Police, public security organ, procuratorate, customs, frontier forces, etc.

D. The reason of choosing DE2-70 board, Nios-II & SOPC

1) The Altera DE2-70 FPGA multimedia development platform is strong in multimedia process, because it is equipped with abundant logic unit, large capacity of memory & various kinds of external chips & peripheral interfaces, supporting multi-media process like video, audio, images, etc. Also, it has the advantages of low power consumption & high integration level.

2) When complex computation is involved, building a SOPC (system on programmable chip) can make the system more efFigient & flexible than describing the hardware architecture purely using Verilog HDL.

3) The Altera's Nios-II embedded processor supports the C language. So a wide range of IP cores can be transplanted, making it easier, more convenient & efFigient to realize the functions.

4) As the Altera devices have already become a mainstream, there are abundant materials & references. Furthermore, Altera also provides excellent technical support. It is so convenient to seek helps.

II. FUNCTION DESCRIPTION

A. Video Surveillance:

Video is filmed by a camera, processed by the FPGA system and finally displayed on the LCD display in VGA mode.

B. Face Detection:

Human faces occur in the video will be detected, intercepted and cached into the SDRAM.

C. Image Compression:

In the SOPC with a transplanted μC/OS-II system, μC/FS system and a Nios-II processor, the cached face image will be read and encoded into JPEG format.

D. Face Image Transmission:

The JPEG images will then be transmitted to the PC in use of MMS through the CDMA modem.

E. Face Image Reception:

The C/C++ program in the PC, which is developed by Visual Studio 2010, can realize the function of MMS reception & JPEG images extraction.

F. Face Recognition:

The program will then analyse the face images based on the PCA (principal component analysis) algorithm.

III. RESOURCES

A. Hardware:

Altera DE2-70 Board, Analogical Camera, CDMA Modem×2, PC, etc.

B. Software:

Quartus II 11.0 sp1, SOPC Builder 11.0 sp1, Nios II IDE 11.0 sp1, ModelSim SE 10.0c, Visual Studio 2010, etc.

IV. ARCHITECTURE

A. Hardware Architecture

The System is divided into two parts—the first part & the second part. The first part is based on FPGA, connected with some peripherals. The second part is based on PC programs and peripherals. In the first part, Verilog HDL was used to describe the hardware modules and Nios-II processor was added into the SOPC. The external chips & peripherals connected to the FPGA system for communication & data transmission are as follows: analogical camera, TV decoder--ADV7180, two 32MB SDRAMs, VGA video DAC—ADV7123, LCD display & CDMA modem. In the second part—PC, the main peripheral is another CDMA modem. The block diagram of system's hardware architecture is shown in **Fig.1**. Illustrations of the signals from No.1 to No.20 are as follows:

(1). Interlaced analogical video input;
(2). NTSC digital video, video control signals, such as horizontal & vertical synchronous signals;
(3). I^2C configuration for NTSC video output;
(4). Interlaced YUV4:2:2 video & relative control signals;
(5). Control signals for reading and writing SDRAM1, YUV4:2:2 frames of even & odd fields;
(6). YUV4:2:2 frames of odd fields;
(7). YUV4:2:2 frames of even fields;
(8). Progressive YUV4:2:2 video;
(9). Progressive YUV4:4:4 video;
(10). Progressive RGB video;
(11). top-left corner coordinate of face area and relative control signals;
(12). Feedbacks for frame buffer, deinterlacing and synchronism control;
(13). Progressive RGB video, VGA display control signals;
(14). VGA video output;
(15). RGB face images;

(16). Control signals for writing SDRAM2, RGB face images;
(17). RGB face images, Control signals for reading SDRAM2;
(18). Feedback from Nios-II;
(19). Face images JPEG files & control signals for MMS;
(20). MMS (enveloping JPEG face images);
(21). Received JPEG files;

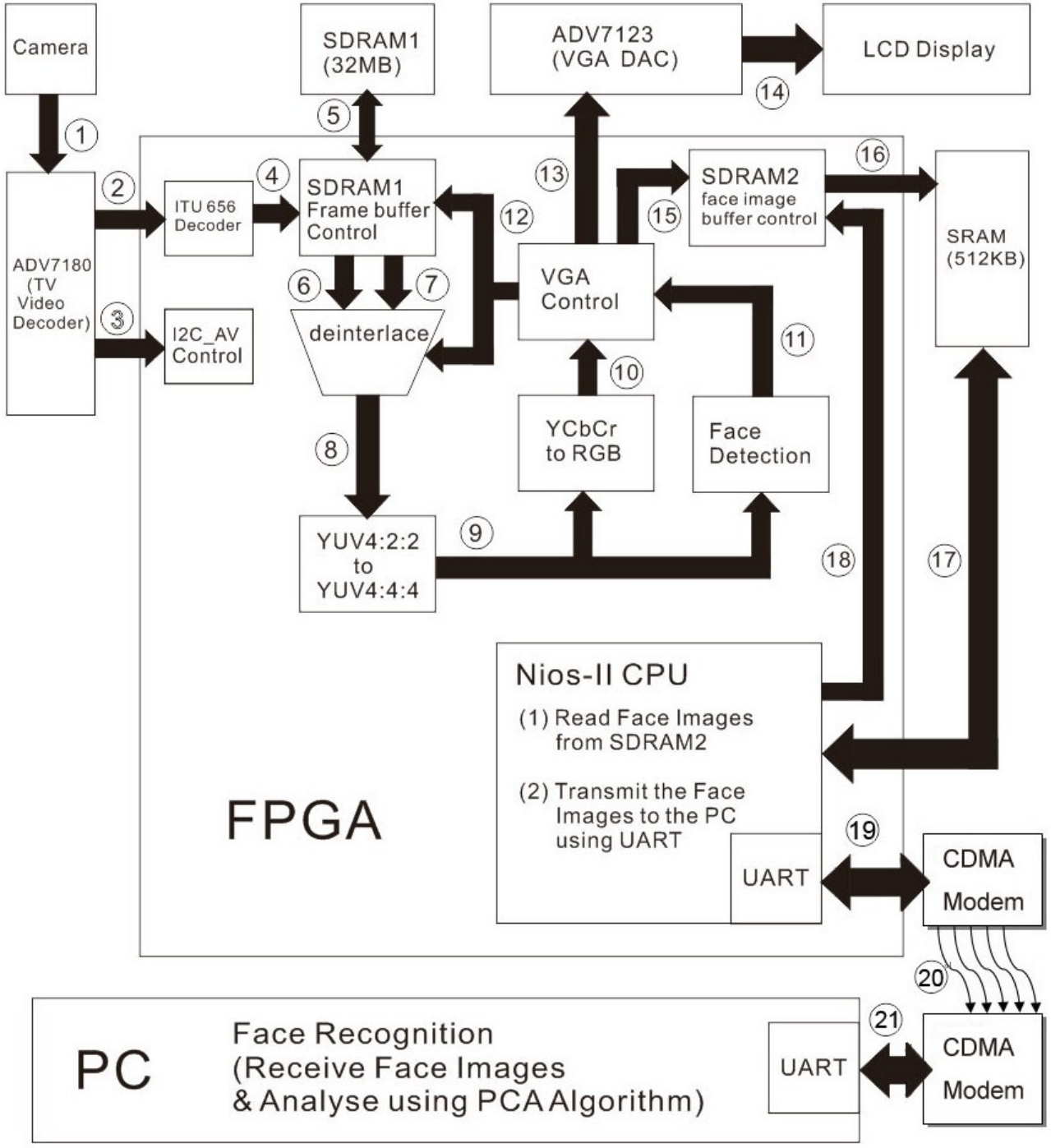

Figure.1 Hardware Architecture Diagram

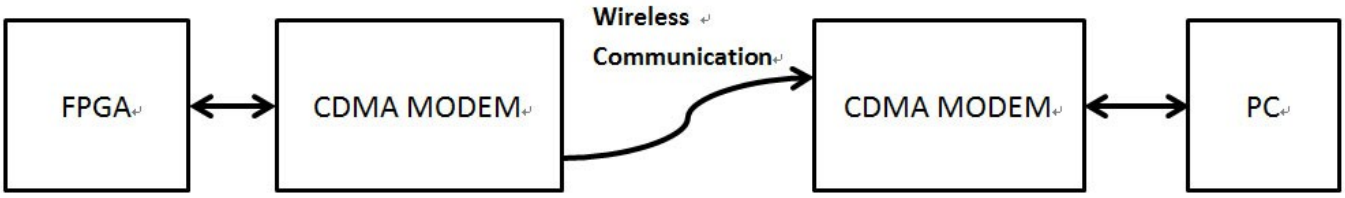

Figure.2 Wireless Communication Diagram

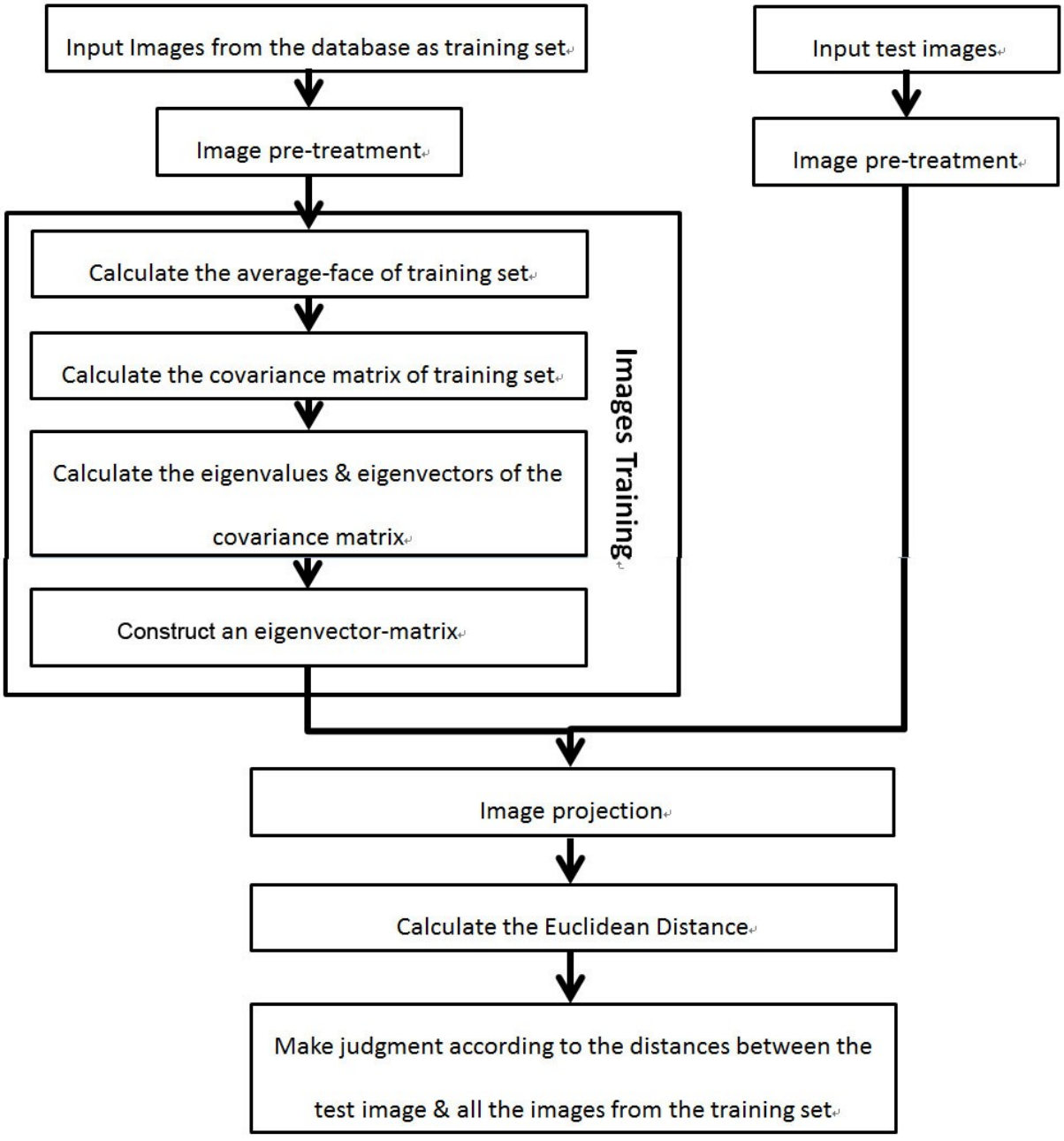

Figure.3 Flow Diagram of PCA program

B. Software Design

We transplant the μC/OS-II & μC/FS into the SOPC and added a Nios-II processor. 3 main tasks have been set:

Task1:

Reading images from the SDRAM2, JPEG encoding & caching;

Task2:

Receiving Data from UART (connected to CDMA modem);

Task3:

Data transmission through UART (connected to CDMA modem) after certain process;

V. DESIGN METHODS

A. Video Capture, Format Conversion & Display

Video capture, format conversion & display are all based on the original demonstration from the DE2-70 CD. A general description about the stages of video process is as follow:

1) Interlaced analogical video are captured by an analogical camera and then input to the ADV7180. After decoding into NTSC by ADV7180, the NTSC video will go through the ITU656 decoder and decoded into YUV4:2:2 digital video;

2) After that, the video will sequentially be deinterlaced, converted from YUV4:2:2 to YUV4:4:4 and transmitted both to face-detecting module & "YCbCr-RGB" module. The "YCbCr-RGB" module will convert the YUV4:4:4 video into RGB video, & transmit them to the VGA controller;

3) The VGA Controller will transmit the RGB video & the relative control signals to ADV7123. Finally, VGA video is output & transmitted to the LCD display.

B. Face Detection

It is known that PCA algorithm requires large amount of computation. In order to reduce the redundant messages & computation as well as a bit of background interference, we built a face detection module. This module can detect human faces from the video, then give out the regional message of the face area and cut out the face image from the original frame at speciFig intervals. Finally, the face images will be encoded & transmitted to the terminal—PC to be analysed.

As shown in **Fig.1**, YUV4:4:4 video was input to the face detection module & skin detection method is used afterwards. When this module operates, it can search out the speciFig area, within which there are maximum quantities of skin-colour-pixels, and it is considered most likely to include human face. Details of how the face detection module functions will be introduced as follows:

1) Each frame will be divided up into a grid with every single 32*32(pixel) square area as a unit. There are 15*20=300 units in each frame because each of them has the resolution of 640*480(pixels). Therefore, each frame can be transformed into a grid matrix with 15 rows & 20 columns.

2) 20 counters are set up corresponding to the 20 grids in one row of grid matrix and a memory of 20bits, 15 units is set up as flags of each grid. Each unit (20 bits) corresponds to each row. Then, in the current row, quantities of pixels within the range of skin-color in each grid will be counted, and the corresponding flags will be set to "1" or "0", as a result. If there are more than a half of pixels are within the range of skin colour (skin-colour pixel) in one grid, indicating that this unit is more likely to include human skin, the relevant flag will be set as "1"—positive. Otherwise, the flag will remain "0"—negative.

3) Statistics will be in synchronous procession with the progress of display. As soon as pixels-counting for one row of grid is finished and the relevant flags are triggered, the 20 counters will be reset so that counting on a new row can be started immediately.

4) When the statistics of one frame is finished, a 15*20 flag matrix is obtained. Then, a full range searching will starts, using a 5*5(flags) square for comparison. The searching will find out the speciFig location, where there are maximum quantities of positive flags with the area of 5*5(flags). Such area will be considered most likely to include human face.

5) Finally, the coordinate of the face-area obtained from the previous stage will be sent to SDRAM2-face buffer controller. Then the face area images can be cut out and cached into the SDRAM2.

C. Face Images JPEG Encoding

JPEG formatting are used to compress each face image in order to save the room of storage, improve system's performance & efFigiency. The reasons why JPEG is chosen are described as follows:

1) Considering the practical situation, the face-recognition system will be deployed & supervise a certain place constantly, there are huge number of images being captured, converted, transmitted & analysed in one day, or even in one hour. Thus the requirement of system's computation ability is extremely high. Since certain kinds of hardware are chosen, its performance, computation ability & data transmission speed have been fixed, the best way to improve the system's performance & efFigiency is to improve the algorithm & software. We should always transmit the most messages in use of the least amount of resources.

2) While the image qualities are at the same level, JPEG files have a higher compression ratio than other image formats. In other words, JPEG encoding can retain a good visual effect while making use of the least storage room.

3) MMS supports JPEG images.

In order to use JPEG encoding, the μC/OS-II & μC/FS are transplanted into the SOPC. A task is set to activate JPEG encoding. Immediately when a new face image has been stored into SDRAM2, a flag will be set as "1" & it will trigger the task. Nios-II will then read the image, encode it

into JPEG format, cache, and set the flag back to "0". After that, when the JPEG image is available for transmission, another task will transmit the image, while another new image will be stored into the SDRAM2 in the same time, covering the old one, and the flag will again been set as positive. The writing & reading process of the SDRAM2 must be carefully designed so that they can keep in synchronism.

D. Data Transmission & Reception

Two CDMA modems are used to establish communication between the FPGA system & the PC program, as shown in **Fig.2**.

The transmission MMS (enveloping a face image) can be divided into two parts. The first part is the JPEG data transmission based on XMODEM Protocol. The second part is the customized AT instructions to defined the functions of wireless module. Some kinds of AT instructions mainly used are as follows:

(1). AT$FDEL="filename"
----delete the file from the module
(2). AT$FUPL="filename"
----copy the file from the PC into the module
(3). AT$MMSW=0
----reset all MMS parameters
(4). AT$MMSW=1,1
----set destination number or e-mail address
(5). AT$MMSW=4,[]
----set MMS title
(6). AT$MMSW=5,[],[]
----set MMS text
(7). AT$MMSW=6,[],[<filename>]
----set MMS accessories
(8). AT$MMSWEND
----send MMS

In the FPGA system, apart from task 1, two more tasks are set up, number as task 2 & 3. Task 2 is set to deal with the data received from the UART (connected to the CDMA modem). Task 3 is set to cache data & transmit them to the CDMA modem through UART.

In the multithread PC program written by C/C++, the JPEG image files will be extracted from the received MMS, & then send to the face-recognition module for PCA.

E. Algorithm of Face-Recognition
(based on PCA—Principle Component Analysis)

The core idea of PCA is: project the image into the eigenvector-space in order to eliminate the trivial messages while retaining the main messages, so that the differences between the two images could be find out in a more efFigient, resource-saving way. In this case, a complete PCA progress includes 7 main steps. The detailed flow diagram of the progress is shown in **Fig.3**. The 7 steps are listed in detail as follows:

(1). Calculate the average face image of the face database;
(2). Calculate the covariance matrix of the database (average face is necessary);
(3). Calculate the eigenvalues & eigenvectors of the covariance matrix;
(4). Set M to be a natural number. Use the eigenvectors relevant to the M maximum eigenvalues to construct an eigenvector-matrix;
(5). Project every face image (including faces from the database & the faces captured to be analysed) into the eigenvector-matrix. The projection of each face is also a matrix (image) called eigenface, which indicates the main characteristics of the original face image.
(6). Calculate the Euclidean Distance between the eigenface being analyzed and the eigenface obtained from the database. The Euclidean Distance indicates the similarity between the two faces.
(7). Make judgment on who the face being analysed is belong to, according to the magnitude of the Euclidean Distance. If the distance is larger than the threshold, the two faces will be considered belonging to the same person; otherwise, belonging to different people.

The core-codes of the PCA program are as follows:

```
//********************************************************************************
// (1) Training:
// calculate the average face & construct the eigenvector-matrix
void doPCA()
{   int i;
    CvTermCriteria calcLimit;
    CvSize faceImgSize;
    // adjust the size of the image
    faceImgSize.width  = FACE_WIDTH;
    faceImgSize.height = FACE_HEIGHT;
        for(i=0; i<faceImgArr.size()-1; i++) {
        eigenVectArr.push_back(cvCreateImage(faceImgSize, IPL_DEPTH_32F, 1));
    }
```

```
    // calculate the eigenvalues
    eigenValMat = cvCreateMat( 1, eigenVectArr.size(), CV_32FC1 );
    //a series of face images
    pAvgTrainImg = cvCreateImage(faceImgSize, IPL_DEPTH_32F, 1);
    // define the condition of ending iteration
    calcLimit = cvTermCriteria( CV_TERMCRIT_ITER, eigenVectArr.size(), 1);
    // calculate the average face, eigenvalues & eigenvectors of the covariance matrix
    cvCalcEigenObjects(
        faceImgArr.size(),
        (void*)faceImgArr.data(),
        (void*)eigenVectArr.data(),
        CV_EIGOBJ_NO_CALLBACK,
        0,
        0,
        &calcLimit,
        pAvgTrainImg,
        eigenValMat->data.fl);
    cvNormalize(eigenValMat, eigenValMat, 1, 0, CV_L1, 0);
}
//*********************************************************************************
// (2) Face-recognition:
//normalize the test face image
        if(faceImage)
            processedFaceImage = equalizeImage(faceImage);

        // face-recognition
        if(processedFaceImage)
            if (eigenVectArr.size() > 0) {
                // project the face image into the eigenvector-matrix.
                cvEigenDecomposite(
                    processedFaceImage,
                    eigenVectArr.size(),
                    eigenVectArr.data(),
                    0, 0,
                    pAvgTrainImg,
                    projectedTestFace);

                // similarity comparison
                iNearest = findNearestNeighbor(projectedTestFace, &confidence, noseRect, mouthRect,
leftEyeRect, rightEyeRect);
                printf("Most likely person in camera: '%s' (confidence=%f.\n",
faceToPerson.at(iNearest)->getName().c_str(), confidence);
```

(a) (b)

(c) (d)

Fig.4-(a),(b),(c),(d) Results of face detection

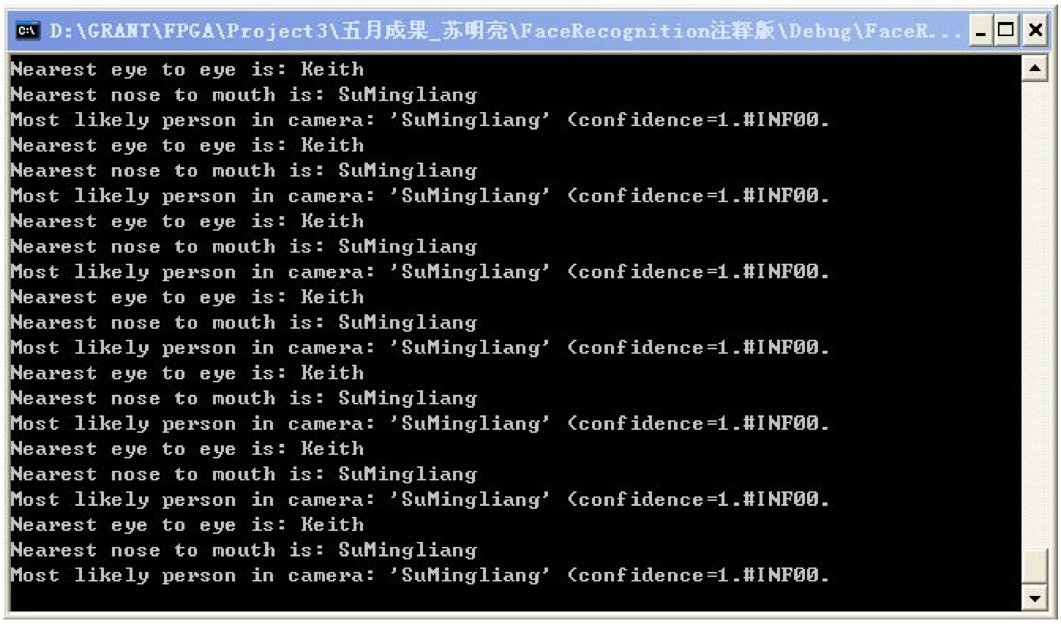

Fig.5 Results of Face Recognition

VI. EXPERIMENTAL RESULTS

A. The 1st stage (the first part):

The operation results of the face detection module are shown in **Fig.4**. Human faces can be detected from the input video and a black square frame will frame the human face region. Therefore, the image inside the frame will be intercepted, encoding & transmitted. However, improvement on the face-detection accuracy is still needed.

B. The 2nd Stage (the second part):

The result of the second stage is shown in **Fig.5**. The face image on the right side is the one to be analysed, while the output window on the left will give out the conclusion.

VII. SUMMARY

Up to now, the *automatic video security system* has achieved the primary goal of face detection, face recognition, and wired communication between PC and FPGA. But the accuracy of face detection & recognition still needs further improvement. The communication still needs perfection, too.

Through this competition, we learned a lot of knowledge about video & image process methods, Verilog HDL, SOPC, μC/OS-II operating system, μC/FS file-management system, FPGA function simulation and timing simulation, SDRAM control, CDMA Modem control, MMS transmission & reception, PCA algorithm and so on. Valuable experience has been accumulated, too.

Compared with the traditional method of designing digital circuit, we clearly realize the distinct advantages of EDA software, Verilog HDL & SOPC.

When dealing with complex problems like PCA and FPGA itself isn't capable for complex computation, it can still control & makes use of the external system with strong computation ability through wireless communication. Therefore, it is still able to accomplish the task without losing the advantages of FPGA.

This design makes use of the power of the server with strong computation ability and the existing wireless communication network. It also makes use of the advantages of FPGA such as high speed, high integration level, low power consumption & great convenience.

By combining the FPGA, PC and wireless communication network, we introduce a possible pattern for real-time video face recognition. As long as the algorithm (accuracy & consumption) & hardware's performance are improved, such kind of system may exert its power to assist the police combating against evil & crime in order to protect the people.

In the end, we would like to express our gratitude to the Altera Corporation who provides a good platform for FPGA project-building exercise & gave us a chance to practise designing & building FPGA system. We would also like to thank the teachers and all the seniors & classmates that had ever helped us! Of course, as a team, all members devoted themselves into the project & work hard. We should finally thanks for ourselves & the effort that we ever made.

ACKNOWLEDGMENT

This work was supported in part by the Strategic emerging industry key technology special project of Guangdong Province (2011168014 and 2011912004).

REFERENCES

[1] Jin-Ming Wang made, Digital System Design and Verilog HDL (The Fourth Edition), Electricity Industry Press

[2] The Science Major in Nan-Jing University adapted, Terasic DE2-70 Development Board FPGA/SOPC Entry Level Experiment Guide Book

[3] Rui-Zhen Liu, Shi-Qi Yu made, OpenCV 刘瑞祯，于仕琪 编著，OpenCV Course (Entry Level), BeiHang University Press

[4] Juan-Su Zen made, The Software Design of the System of Wireless Infrared Anti-theft Alarm Basing on MMS, HeBei Industry University 2007

[5] Jin Duan made, Automatic Human Face Recognition, Science Press

[6] Jin-Shu Chen, He-Xin Chen, Ai-Jun Sang made, The Technology Summary on Human Face Recognition on Computer, Jilin University Newspaper(Information and Science Edition) Volume twenty-first Suppl. 2003,8

[7] Xue-Yin Yu, Kuo-Huang Cao, Ben-Sheng Chen made, Human Face Recognition Based on Eigen Face and Its Realization, HeBei Industry Science, Volume twenty-fifth, 2009,9

[8] Ren-Yan Zhou, Bin Shang made, Visual C++ Serial Communication Development and Programming Practice, Electricity Industry Press, 2009

[9] Min Tian, Yao Zheng , Jiang-Quan Li etc. made, Visual C++ Data Collection and Serial Communication Control Application of Actual Combat, People Post and Telecommunications Press, 2010