A Low Complexity Real Time Face Tracking System with Fuzzy Controller

Chih-Yung Chen, Huang-Chu Huang, and Rey-Chue Hwang

Abstract

This paper proposes a new face tracking system, which emphasizes on decreasing the face tracking computation complexity and manipulation procedure in face detection for a real time application. Face detection and camera servo motor control are two main parts of the face tracking system proposed. Face detection approach is based on image processing techniques. The techniques include discrete wavelet transformation (DWT) image data reducing, motion detection, edge detection, fast modified Hough transformation and face color detection by using probabilistic neural network (PNN). A two-dimensional fuzzy controller is used for the execution of camera servo motor control. It could make the camera fast and exactly detect the human face on its center. From the experimental results, such a face tracking system has shown its high potential in real applications.

Keywords: Face tracking, face detection, fuzzy controller.

1. Introduction

Face detection and recognition are the popular topics in computer vision research area. The application scopes include security guard management, transaction authentication, residential monitoring etc. Generally, the discovery of human faces, and the number of human faces estimation in an image, etc. are defined as the face detection problem. Further, if the objective of camera was used to trace the human face in a captured scene, then such a research will be treated as a kind of face tracking problem. Basically, the first procedure of face tracking is to capture the face image by camera, e.g. charge-coupled device camera (CCD), then the image processing algorithms will be continuously used to detect where the face is. Finally, the camera will moves to an appropriate position so that the camera lens can aim at the face location.

Corresponding Author: Professor Rey-Chue Hwang is with the Electrical Engineering Department of I-Shou University, Kaohsiung County, Taiwan, 840.
E-mail: rchwang@isu.edu.tw

This paper mainly focuses on designing a low-complexity and high efficiency face tracking system. In the face detection process we proposed, all techniques used are based on image processing. Duo to shrink the image size and make the signal's information can be effectively processed, discrete wavelet transformation (DWT) is used to reduce the size of image data. After the image reducing transformation, two methods, shape detection and color detection, are then used to detect the face location. In shape detection, the motion detection firstly finds out the changed region, i.e. the interesting region, and then the morphology edge detection detects the edge information. The fast ellipse Hough transformation (FEHT) is used to find the shape of ellipse and the center of ellipse. Finally, PNN model is used for skin color recognition. It will classify the face color pixels in the regions detected.

In the tracking process, the camera will be moved to aim the human face by two axes servo motors which are controlled by a two-dimensional fuzzy controller. Further, the servo motors are droved by pulse width modulation (PWM) controller which is designed by 8051. The detailed flow chart of proposed approach is shown Figure 1.

2. Face Detection

In the researches in face detection, the main problem is the execution speed of detection system could not achieve the requirements for a real time application. To overcome such a problem, the system developed uses several effective detecting methods that will be described in the following paragraphs. The whole system can execute 10-30 frames per second based on the size of image and the processing speed can also meet the requirements of the real application asked for.

A. *Discrete wavelet transformation*

DWT techniques have been widely used in the signal processing problems. The Lifting 5/3 [1] is the simplest one of DWT and it can be implemented by hardware easily. Due to the image size reducing, Lifting 5/3 is selected in the proposed scheme with respect to its low complexity. The Lifting 5/3 DWT diagram is presented

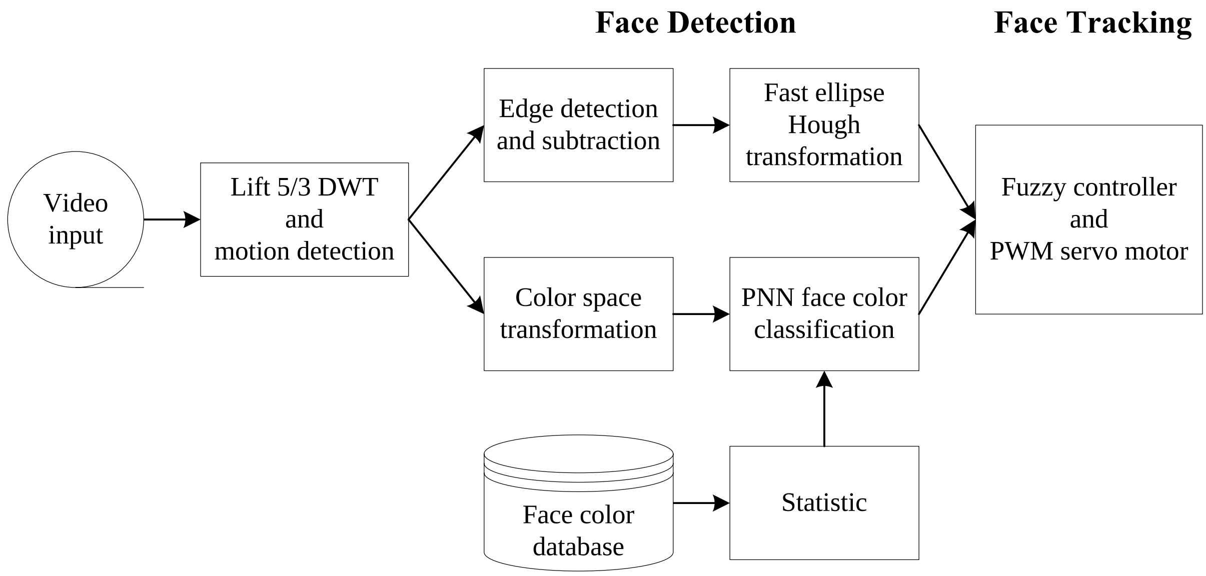


Figure 1. The flow chart of proposed approach.

as Figure 2. Generally, Lifting 5/3 DWT has several advantages, including low computation, less memory requirement, easily forward and inverse transformations, and suitable for integer-to-integer transformation. Basically, the procedure of Lifting 5/3 can be divided into two steps:

1. Splitting step

In this step, a raw data R with length n can be divided into two parts, the odd points r_{2i+1} and even points r_{2i}. Here, they are denoted as S_i^0 and D_i^0, respectively. Their mathematical expressions are then given by

$$S_i^0 = r_{2i+1}, \quad D_i^0 = r_{2i}, \quad \forall (r_{2i+1}, r_{2i}) \in R \ \text{, and} \ i = 0,1,2,...\frac{n}{2} \tag{1}$$

2. Lifting step

The high-pass elements D_i^1 and low-pass elements S_i^1 can be calculated by

$$D_i^1 = D_i^0 - \frac{(S_i^0 + S_{i+1}^0)}{2}, \ i = 0,1,2,...\frac{n}{2} \tag{2}$$

$$S_i^1 = S_i^0 + \frac{(D_{i-1}^1 + D_i^1)}{4}. \quad i = 0,1,2,...\frac{n}{2} \tag{3}$$

In image processing, the captured two-dimensional image can be firstly transformed by using horizontal and vertical Lifting transformations. The image then can be divided into four bands, i.e. { LL, LH, HL, HH }, where L and H stand the low-band and the high-band, respectively. The LL band is the quarter image of original image as shown in Figure 3. The image was reduced and filtered by DWT, which could decrease the computation complexity and eliminate noises.

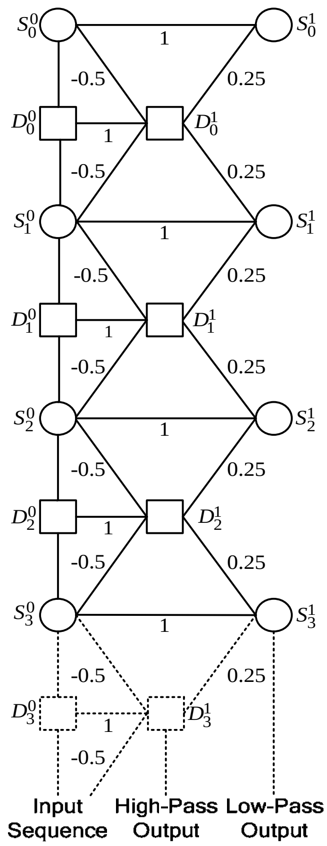


Figure 2. Lifting 5/3 DWT.

Figure 3. DWT transformed image.

B. *Motion detection*

In motion detection, the proposed approach uses the gray-level image subtraction for image processing. Suppose a captured image sequence $\mathbf{I}$ with length l. I_t is one image belongs to $\mathbf{I}$ with n by m image size, then

$$\forall I_t \in \mathbf{I} \ \text{ and } \ \forall P_{i,j}^t \in I_t \tag{4}$$

where i and j stand the pixel P's row and column indexes of image I_t.

The gray-level image subtraction of two continuous images I_t and I_{t+1} can be defined as

$$\forall M_{i.j} \in \mathbf{M} \ \text{ and } \ M_{i.j} = \begin{cases} 1 & , \left| P_{i,j}^t - P_{i,j}^{t+1} \right| > \lambda \\ 0 & , \text{otherwise.} \end{cases} \tag{5}$$

where $\mathbf{M}$ is a binary two dimensional vector and λ is a constant threshold which can be set by the designer. In this study, $\lambda = 50$ is taken.

Usually, the process of real time motion detection will be significantly affected by the high frequency noise. In order to avoid the effect caused by high-frequency noises, the DWT transformed LL image is mainly considered and taken for the signal processing due to it has less high-frequency signals in comparison with other band's image. Hence, the performed $\mathbf{M}$ can be used to find the changed pixels, i.e. "1" of equation (5), and can be used as the judgment criterion for the object's movement desired to track.

C. *Morphology edge detection*

Except the face color information, face sketch is an alternative way to detect face location. In this study, morphology edge detection is used to find the edge of captured image.

The idea of mathematical morphology was initialized by G. Matheron in year 1975 [2]. Serra then proposed the morphology theory which includes set theory, geometry and algebra in year 1982 [3]. In our study, only grey scale operations are discussed. In addition, only dilation is used and reported. Moreover, morphology image processing mainly operates the structuring element which is a set of neighborhood pixels. The proper structuring element assignment will greatly improve the whole processing result. For describing the morphology implementation, suppose we have an image X and a pre-defined two-dimensional structuring element U. In X, each pixel is denoted as $x_{i,j}$, where i and j are pixel's row and column indexes. The size of U is $n_s \times n_s$. The operations of edge detection can be described as follows:

1. Dilation

In image processing, dilation operation can extend the boundary of an object by removing the low valued regions. After the dilation operation, new pixel $x_{i,j}^D$ can be obtained by equation (6).

$$x_{i,j}^D = x_{i,j} \oplus U = \max\left\{ x_{i+(a-n_s/2),\, j+(b-n_s/2)} + u_{a,b} \right\} \tag{6}$$

where, $u_{a,b} \in U$ and $a, b = 0, 1, 2 \ldots n_s$.

2. Edge detection

It is well known that the edge detector of image segmentation usually has a huge computational complexity. In our operation, only one dilation and one subtraction are operated for each pixel. The morphology edge detection then can be simply and efficiently calculated by the following equation.

$$x_{i,j}^{Edge} = x_{i,j}^D - x_{i,j} \tag{7}$$

The edge detection result is shown in Figure 4. Besides, in order to find the motion object's contour, the gray-level image subtraction is repeated in this process. Suppose two continuous images have X_t^{Edge} and X_{t+1}^{Edge} edge results, respectively. Equation (5) can be taken for calculating the subtraction result, i.e. contour of two images.

Figure 4. Edge detection result.

D. *Fast ellipse Hough transformation*

Normally, the shape of human head is near an elliptic. After the motion object contour was found by using the methods described in the previous paragraphs, an elliptical shape searching algorithm is helpful for face detection. Hough transformation is a well-known shape detection technique of image processing. The face tracking used Hough transformation to detect the elliptical head can be referred in article [4]. However, the traditional Hough transformation usually has the problem of computation complexity. In the approach we proposed, a fast ellipse Hough transformation (FEHT) is used to meet the requirements in the real time application. The detailed description of FEHT was reported in article [5]. For improving the efficiency of FEHT, the rate of x-axis (short axis) and y-axis (long axis) radiuses of ellipse is set to 1:1.2~1.4. In the system we developed, any ellipse whose rate is over the value we set will not be processed. That means too small or too big ellipse will be dropped. Additionally, FEHT only search the ellipse in the region of motion detected. Such strategy can effectively reduce the whole processing complexity.

E. *Skin color detection*

Skin color information is a very important function for the face detection [6]. It is a common way used in the human face detection. In our approach, a probabilistic neural network (PNN) is taken as the human skin color classifier. PNN was proposed by Specht [7], is a type of supervised neural model. The architecture of PNN is shown in Figure 5. It is a direct neural network implementation of the Parzen windows method and Bayes decision rule. One input color pixel of a captured image taken by the camera can be transformed into a one-dimensional vector $\mathbf{g}$, and denoted by

$$\mathbf{g} = \{g_1, g_2, \dots g_m\} \tag{8}$$

where m is the length of color pixel, e.g. the length of RGB color space is 3.

In the system we developed, the face color database was constructed firstly based on many photo shoots. The supposed face color classes $\mathbf{c}$ with r categories is then built, it can be represented as

$$\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_r \quad \text{and} \quad \mathbf{c}_j = \{y_{j1}, y_{j2}, \dots, y_{jm}\} \tag{9}$$

where j is the index number of classes $\mathbf{c}$ and y is the components of each pixel .

The Parzen probability density function (PDF) is simply obtained by the sum of small multivariate Gaussian distributions centered at each training sample. Therefore, the Parzen PDF estimator can be expressed by

$$f_j(\mathbf{g}) = \frac{1}{(2\pi)^{d/2}\sigma_j^d}\left(\frac{1}{m}\right)\sum_{l=1}^{m}\exp\left[-\frac{(\mathbf{g}-y_{jl})^2}{2\sigma_j^2}\right] \tag{10}$$

where σ_j is the smoothness parameter and d is total number of training patterns.

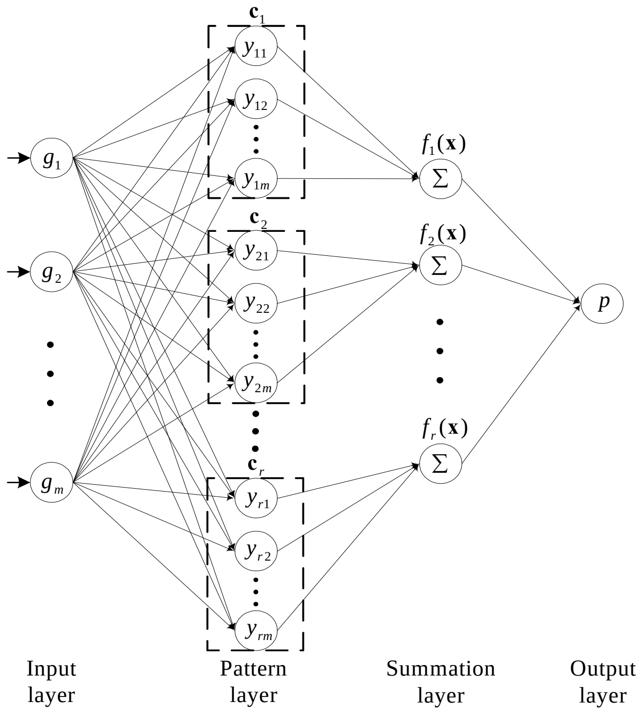


Figure 5. PNN architecture.

After PNN implementation, the maximum probability p of an input pixel can be obtained. If p is larger than the threshold value we defined, then the pixel will be determined as the face color pixel. Since the number of supposed face color set in this study is 50, so that the classes of pattern layer and the categories of summation layer of PNN are all set to be 50. As stated above, only the captured image pixels in the region detected by motion detection are used as the input data for PNN classifier.

The four algorithms mentioned above are combined to be the complete face detection procedure and used to detect the face position for the tracking system. The flow chart of the proposed face detection diagram is shown in Figure 6.

3. Active Fuzzy Servo Motor Controller

In the hardware approach of the system we proposed, a two-dimensional tilt platform is designed. It can make the 180-degree movements from left to right and top to

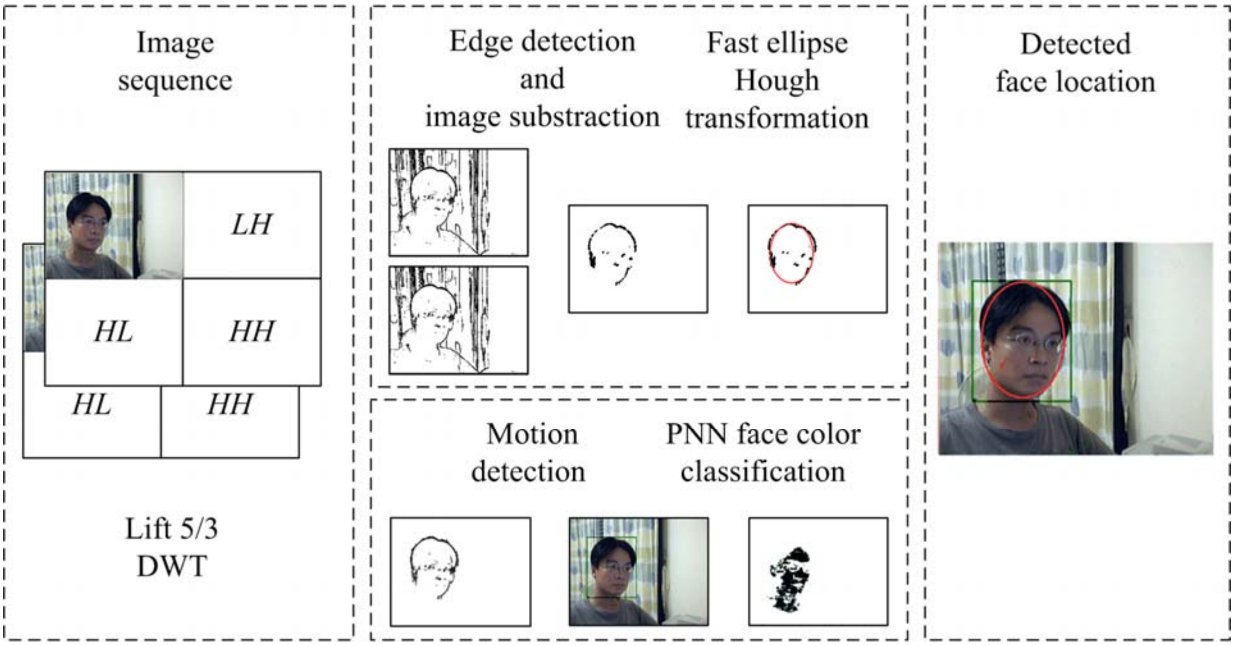


Figure 6. The proposed face detection method.

down, respectively. The panel appearance is shown in Figure 7, which contains a camera, bearings and two servo motors.

Figure 7. The two-dimensional tilt platform.

The control method for the servos motor is pulse width modulation (PWM). In order to track the face location precisely, a two-dimensional fuzzy controller is used to control the motors. Generally, once the face location is found by face detection method, the move servo motors should be able to make camera aim at the center of face region quickly and then follow the target. Reasonably, the one step method is a perfect way should be used. However, since the parameters such as the angle and distance between human and camera, the speed of servo motor and relationships between the rotation angle and captured image etc., are hardly defined. A perfect one step tracking control method is hardly and impossibly developed. Contrary, an improper one step control method will make camera easily miss the target during its tracking process. Therefore, for efficient handling the uncertainties of environments encountered, fuzzy controller is a suitable way to be used in our system. The detailed descriptions of the controller we developed are described as follows.

A. *PWM control*

In our controller, the pulse width range of servo motor used is from 0.7 milliseconds to 2.3 milliseconds. It can rotate the servo motor form 0 degree to 180 degree. The duty cycle of serve motor is 20 milliseconds. Its pulse characteristic diagram is shown in Figure 8. Beside, the PWM module is designed as firmware which can be easily embedded on a microprocessor.

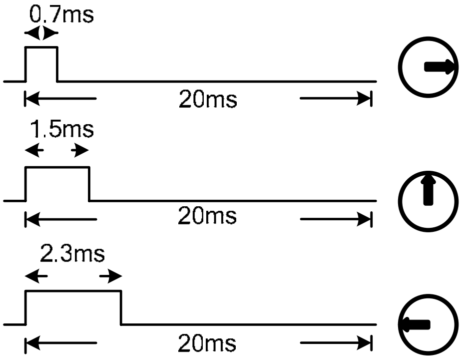


Figure 8. Servo motor characteristics.

B. *Fuzzy control*

In our fuzzy control system, the increments of *PWM* will be determined by a two-dimensional fuzzy controller. The fuzzy mechanism is basically described as follows [4, 8].

Fuzzifier: triangular membership function.
Inference engine: Mamdani max-min inference engine.
Defuzzifier: center average computation.

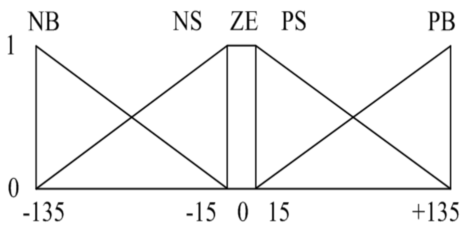
Two parameters are considered to be the fuzzy variables, one is the difference of x-axis between current position (x_p) of face in the image and the image center point (x_c), the other one is the difference of y-axis between current position (y_p) of face in the image and the image center point (y_c). If both variables are denoted as dx and dy , then

$$dx = x_c - x_p , \tag{11}$$

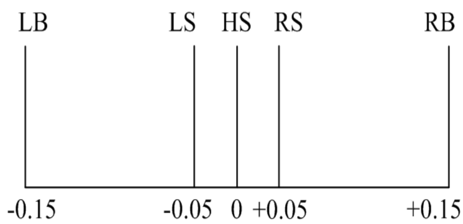
$$dy = y_c - y_p . \tag{12}$$

The fuzzy rules for x-axis controller are:
Rule 1: IF dx = NB, THEN output= LB (left-turn big)
Rule 2: IF dx = NS, THEN output= LS (left-turn small)
Rule 3: IF dx = ZE, THEN output= HS (no change)
Rule 4: IF dx = PS, THEN output= RS (right-turn small)
Rule 5: IF dx = PB, THEN outpu = RB (right-turn big)

The membership functions (MF) of fuzzification and defuzzification of dx are shown in Figure 9(a) and Figure 9(b), respectively.



(a) The MF of fuzzification of dx.

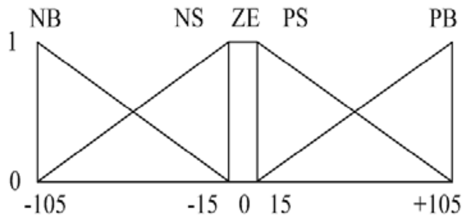


(b) The MF of defuzzification of dx .

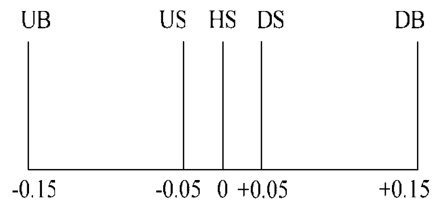
Figure 9. The membership functions of dx .

The fuzzy rules for y-axis controller are:
Rule 1: IF dy =NB, THEN output= DB (down-turn big)
Rule 2: IF dy =NS, THEN output= DS (down-turn small)
Rule 3: IF dy =ZE, THEN output= HS (no change)
Rule 4: IF dy =PS, THEN output= US (up-turn small)
Rule 5: IF dy =PB, THEN output= UB (up-turn big)

The MFs of fuzzification and defuzzification of dy are shown in Figure 10(a) and Figure 10(b), respectively.



(a) The MF of fuzzification of dy .



(b) The MF of defuzzification of dy .

Figure 10. The membership functions of dy .

4. Experiments

In this study, a CCD camera is used for capturing image and its size range can be from 160×120 to 640×480 pixels (RGB 24 bits). The 8051 Microcontroller (MCU) is designed and used as the servo motor controller and the Keil C is used as the firmware developer. The developed MCU module is shown in Figure 11. For speeding up the signal processing, the multithread programming skill is used in software programming of the system. It is able to capture image, execute image processing tasks and send the fuzzy controller outputs to 8051 MCU via RS-232 to control the servo motors at almost the same time. From the experiments under running on a 2.4 GHz computer, the real time tracking frames per second for different size of images are listed in Table 1. According to the results shown, the face tracking system we developed has its potential in real time application very promising. Figure 12 shows some examples of tracking performance by using this system.

Figure 11. PWM MCU control module.

Table 1. Processing frame rates of proposed approach with different resolution.

Image size (pixels)	Frames per second (fps)
160×120	Over 30 fps
320×240	20~25 fps
640×480	10 fps

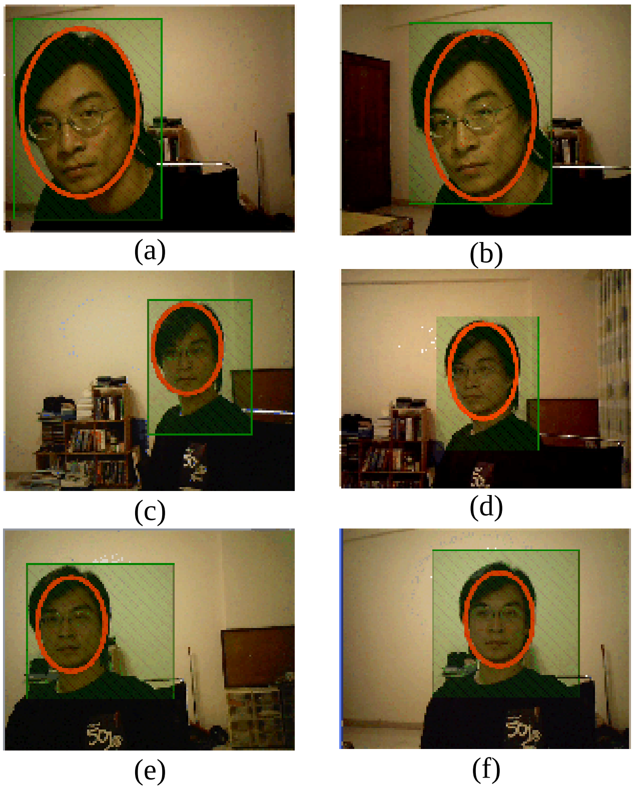


Figure 12. The examples of tracking results by proposed system..

5. Conclusions

In this paper, a new face tracking system was developed and proposed. The main objective of this system is to emphasize on the fast computation and manipulation. Several image processing techniques were adopted in face detection process, including DWT image data reducing, motion detection, edge detection, fast modified Hough transformation and PNN face skin color detection. All techniques are expected to make sure of the system developed can fast process the signals captured.

A two-dimensional fuzzy controller is used for servo motor control of the camera. From the experimental results shown, the simple fuzzy mechanism we constructed could effectively make the camera have a fast response in the human face tracking and detection. Therefore, we conclude that such a face tracking system has shown its high potential in real time applications very promising.

In our future study, how to raise the accuracy of face detection further and make camera have faster response are the works we should put more efforts.

5. Acknowledgment

This work is supported by the National Science Council of Republic of China under contract No. NSC-95-2221-E-214-38.

References

[1] I. Daubechies and W. Sweldens, “Factoring wavelet transforms into lifting schemes,” *J. Fourier Anal. Appl.*, vol.4, pp. 247–269, 1998.

[2] G. Matheron, “Random Sets and Integral Geometry,” John Wiley & Sons, New York, 1975.

[3] J. Serra, “Image Analysis and Mathematical Morphology,” London: Academic, 1982.

[4] S. S. Ghidary, Y. Nakata, T. Takamori, and M. Hattori, “Human detection and localization at indoor environment by home robot,” *2000 IEEE International Conference* on *Systems, Man, and Cybernetics*, vol. 2, pp. 1360-1365, 2000.

[5] N. Guil and E. L.Zapata, “Lower order circle and ellipse Hough transform,” *J. Pattern Recognition*, vol. 30, no. 10, pp. 1729-1744, 1997.

[6] K. Schwerdt and J.L. Crowley, “Robust face tracking using color,” *Fourth IEEE International Conference* on *Automatic Face and Gesture Recognition, 2000.* pp. 90-95, 2000.

[7] D. F. Specht, “Probabilistic neural network for classification, map, or associative memory,” *Proceeding of the IEEE International Conference on Neural Network*, vol.1, pp. 525-532, 1988.

[8] R. C. Hwang, H. C. Huang, Y. J. Chen, and C. C. Kung, “A more effective fuzzy controller,” *2004 IEEE International Conference on Networking, Sensing and Control*, pp. 1394-1399, 2004.

Chih-Yung Chen received his MS degree in computer science from Tamkung University, Taipei, Taiwan, in year 2000. Currently, he is pursuing his Ph.D. degree in the Department of Electrical Engineering, I-Shou University, Kaohsiung, Taiwan. His current research interests include intelligent system, neural network, and image processing.

Huang-Chu Huang received his Ph.D. degree in electrical engineering from National Sun Yet-Sen University, Kaohsiung, Taiwan, in year 2001. Currently, he is an Associate Professor with the Department of Electronic Communication of National Kaohsiung Marine University, Kaohsiung, Taiwan. His research interests include neural network, artificial intelligence and control.

Rey-Chue Hwang received his Ph.D. degree in electrical engineering from Southern Methodist University, Dallas, U.S.A., in year 1993. He is a senior member of IEEE. Currently, he is a Professor with the Electrical Engineering Department of I-Shou University. His research interests include neural network, fuzzy theory, control and signal processing.