Self-adaptive iris image acquisition system

Wenbo Dong, Zhenan Sun, Tieniu Tan, Xianchao Qiu

National Laboratory of Pattern Recognition, Institute of Automation, Academy of Sciences, No.95 Zhongguancun East Road, Beijing, China

ABSTRACT

Iris image acquisition is the fundamental step of the iris recognition, but capturing high-resolution iris images in real-time is very difficult. The most common systems have small capture volume and demand users to fully cooperate with machines, which has become the bottleneck of iris recognition's application. In this paper, we aim at building an active iris image acquiring system which is self-adaptive to users. Two low resolution cameras are co-located in a pan-tilt-unit (PTU), for face and iris image acquisition respectively. Once the face camera detects face region in real-time video, the system controls the PTU to move towards the eye region and automatically zooms, until the iris camera captures an clear iris image for recognition. Compared with other similar works, our contribution is that we use low-resolution cameras, which can transmit image data much faster and are much cheaper than the high-resolution cameras. In the system, we use Haar-like cascaded feature to detect faces and eyes, linear transformation to predict the iris camera's position, and simple heuristic PTU control method to track eyes. A prototype device has been established, and experiments show that our system can automatically capture high-quality iris image in the range of $0.6\,\mathrm{m}\times0.4\,\mathrm{m}\times0.4\,\mathrm{m}$ in average 3 to 5 seconds.

Keywords: biometrics, iris recognition, image acquisition, pan-tilt-zoom camera, face detection, auto zoom

1. INTRODUCTION

Iris recognition is one of the most reliable methods for personal identification, and potential to be used in many mission-critical applications. [1][2] However, the size of iris is small (11 mm), but the number of iris diameter pixels is large (normally 150 pixels required for recognition). Furthermore, iris texture only exhibits in infrared illumination condition, especially for Asian people. So, it is very difficult to capture iris images in practice.

Many commercial products have been developed by some companies such as Iridian[3], OKI[4], Matsushita[5], LG[6],etc.. Most of the products are non-contacting and acquire iris images at a distance. For aiming at eyes at a distance, systems need users to cooperate with the machine actively, such as stare at the camera or move again and again under instructions, but sometimes it is difficult for users to self-locate their positions, especially for beginners. Although those devices are equipped with cold mirrors, monitors, sound indicators, or LED indicators to guide subject's movements, some users still can not perform well. This problem nearly becomes the bottleneck of iris recognition's application.

Recently, many people have advanced automatic iris image acquisition systems based on some new methods.(for example, systems of Sarnoff[7], Mitsubishi[8] and Yongsei University[15]) They used one or two wide-angle cameras to find subject's face, and then another narrow-angle camera with high-resolution (more than 4 mega pixels) to capture the iris image. The iris camera were set up on the pan-tilt unit and controlled to rotate towards eye region . This method is an important reference to us.

Besides this method, Sarnoff corp. has developed the system of "Iris on the move".[9][10] They did not use the movable cameras, but used a camera array made up with many high-resolution cameras (2048×2048 pixels, 15 frames per second). When people walked through a portal equipped with illumination lamps, his eyes would affirmatively appear in one of the cameras immediately.

Both the two methods use high resolution cameras (more than 4 mega pixels),but CCD with more pixels transfers less frame per second. Furthermore, the high resolution with high frame rate is very expensive. We

Further author information: (Send correspondence to Wenbo Dong)
Wenbo Dong: E-mail: wbdong@nlpr.ia.ac.cn, Telephone: 086 010 62632265

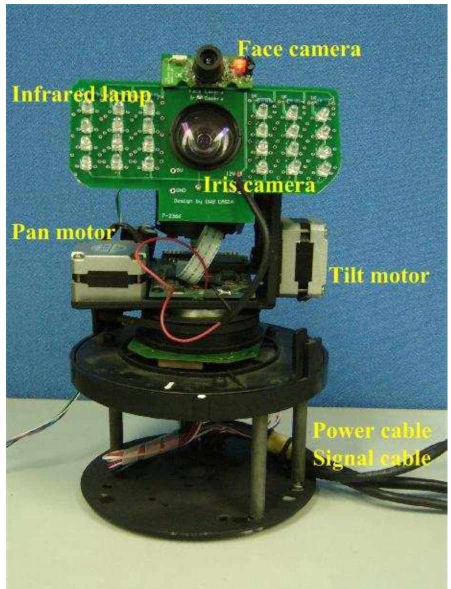

Figure 1. The appearance of the prototype

should look for a tradeoff between high resolution images and devices' price. Is the high resolution cameras the only option or the best option?

In this paper, We use two low resolution cameras (less than 0.5 mega pixels), which are co-located in a PTU, for face and iris image acquisition respectively. Once the face camera detects face region in real-time video, the PTU is controlled to move towards eye region and adjust the lens focus at the same time, until the iris camera captures clear iris image for recognition. Experiments show our system can capture iris image automatically in the volume of about $0.6\,\mathrm{m}\times0.4\,\mathrm{m}\times0.4\,\mathrm{m}$ during average 3-5 seconds(2.7 seconds in the best position).

Our work gave an attempted on non-cooperative iris image acquisition using low resolution cameras, and achieves the same tracking accuracy with the previous techniques. The tracking speed is not very fast, but the image frame rate is fast enough to process images in the real-time, and the design can lower the expense. Moreover, we use Haar-like cascaded feature to detect faces and eyes, linear transformation to predict the iris camera's position, and simple heuristic PTU control method to track eyes. A prototype device has been established. Those are our contributions in this paper.

The paper is structured as follow: The next section describes the overview of the system hardware and software design; section 3 introduces many technique details; while, section 4 and 5 presents the experiment results and conclusions.

2. OVERVIEW OF THE SYSTEM HARDWARE AND SOFTWARE

2.1 Installation

The devices include one wide-angle camera(W-Cam), one narrow-angle camera (N-Cam), infrared lamps, PTU and a computer.

The two cameras and infrared lamps are installed in PTU and rotated together with the PTU. (**Fig. 2**) The wide-angle camera captures face images, and the narrow-angle camera captures eye images. The two cameras are fixed tightly and nearly coaxially. When an eye image appears in a reference point in the W-Cam's center, the eye image will appear in N-Cam's view. The reference point's value is calibrated on installation.

For illumination, we use the infrared LED array with 20 degree emission angle, which is installed on the PTU and rotated with the two cameras together.

Figure. 1 is the appearance of the prototype device we have established.

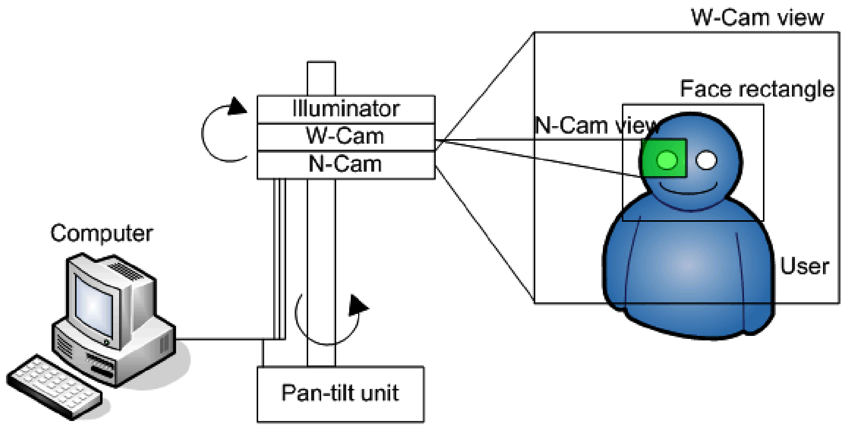

Figure 2. Hardware device installation and connection

2.2 Devices selection

For the wide-angle camera, we select the normal web camera with 640×480 pixels and connect it to computer by USB cable.

For the narrow-angle camera, we select the SONY integrated camera and connect it to computer via a frame grabber card. The camera's CCD has a low-resolution of 768×576 pixels, but can transfer about 30 frames per second. The camera has lens of variable focus from 3 mm to 78 mm and it can zoom near and far automatically. Here we should note that we have re-scaled the camera, otherwise it could not focus the near objects.

For the PTU, we use the type of 100 degree per second both vertically and horizontally. It is controlled by RS-485/RS-232 cable using the standard protocol Pelco-D.

2.3 Software flow

Before software's running, we must calibrate the reference point on the W-Cam.(section 3.3) Then the system software runs as follow: 1. In one software thread, the computer continuously looks for face in the image with the W-Cam, and after finding face, detects eyes in the face rectangle.(section 3.1) 2. If a certain (left or right) eye's position is not on the reference point, the computer moves the pan or tilt motor and rotates cameras to the expected position. 3. In another software thread, computer processes image sequence of the N-Cam and if the eyes is detected, the computer captures the iris image and picks up its feature for recognition.

3. TECHNICAL DETAILS

3.1 Face detection and eyes detection

We use the face detect algorithm based on classifiers cascade that simply use Haar-like features originally selected using AdaBoost learning algorithm.[11] The detection takes less than 100 ms in a 640×480 frame and outputs low false match rate, even if the face is rotated or slanted a little angle.

To detect eyes, we also use the above method based on Haar-like features and AdaBoost algorithm. With this method, we can find the eye positions, specially the pupils, very accurately. But it sometimes recognized eyebrows as eyes when people stand far from the camera. Eyes' standing in the upper part of the face, we can use this contextual information to reduce the search area and consequently the computational time.

We can see the effect of the detection in the following **Fig. 3** and **Fig. 4**.

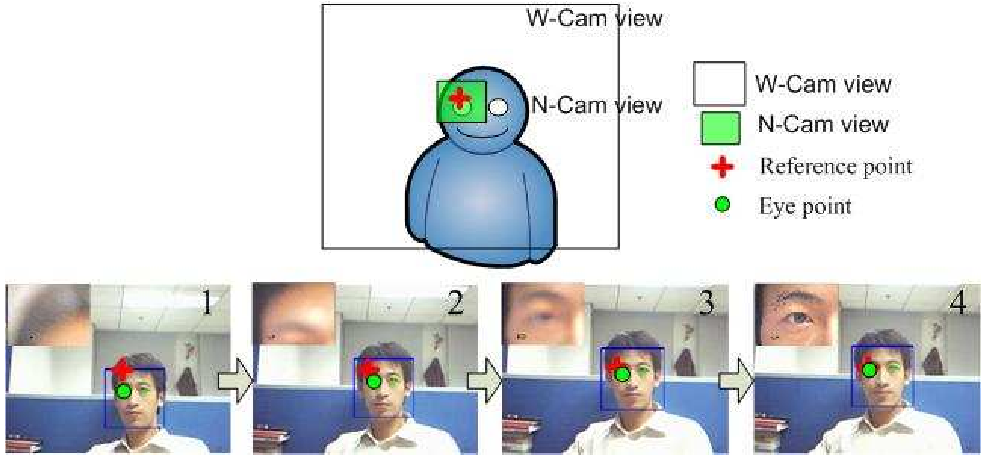

Figure 3. Image mapping of the two cameras

3.2 Imaging scheme based on two cameras

Two cameras are fixed tightly and nearly coaxially, so when an object appears in a reference point of W-Cam (approximately on the center), it will appear in the view of N-Cam at the same time.(**Fig. 3**) So if the eye's position is not on W-Cam center, we can turn the pan-tilt unit, until it arrives the expected position. From **Fig. 3**, we can see an example of the tracking progresses.

Compared to the method of full-Euclidian space calibration using the two cameras, this method is easier to realize, and need no more repeated computation in the software, but has same accuracy with the former in this special context.

3.3 Calibrate the W-Cam's reference point

It should be noticed, the two cameras are not coaxial perfectly, so the eye's real reference point is not on the very center of the W-Cam, but Y pixels under the center. Moreover, Y is not a constant value, but will change with user's distance. (**Fig. 4**)

We can assume that Y has 1-dimension linear relationship with user's distance D, and D is relevant with the face image size in the camera. So Y has 1-dimension linear relationship with the face height H (or face width W).

$$Y = aH + b \tag{1}$$

We will use the linear regression method to calculate the parameter value of a and b. In the **Fig. 4**, when an user is near from the camera, the reference point is Y_1 under W-Cam's center, and the face image height is H_1; when people stand middle, it is Y_2 and H_2; and when people stand far, it is Y_3 and H_3; as such, we can get multiple arrays of (Y, H). Then we can calculate parameters of a and b from the above data.

Let $Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix}$ and $H = \begin{bmatrix} H_1 & 1 \\ H_2 & 1 \\ \dots & \\ H_n & 1 \end{bmatrix}$, for $Y = H \begin{bmatrix} a \\ b \end{bmatrix}$, we get the least square solution

$$\begin{bmatrix} a \\ b \end{bmatrix} = (H^T H)^{-1}(H^T Y) \tag{2}$$

In our system, we get $a = 0.14$ and $b = 10.2$. Though a coarsely estimated equation, it shows good performance in the latter experiments.

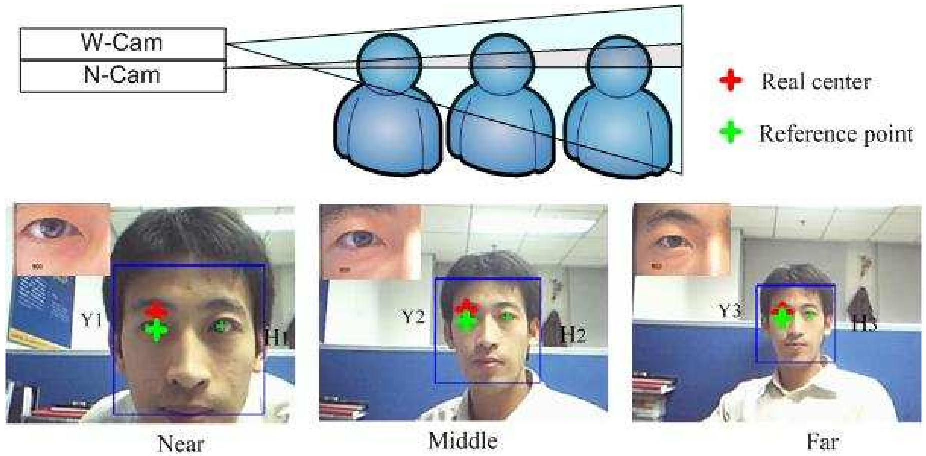

Figure 4. The W-Cam's reference point is not constant

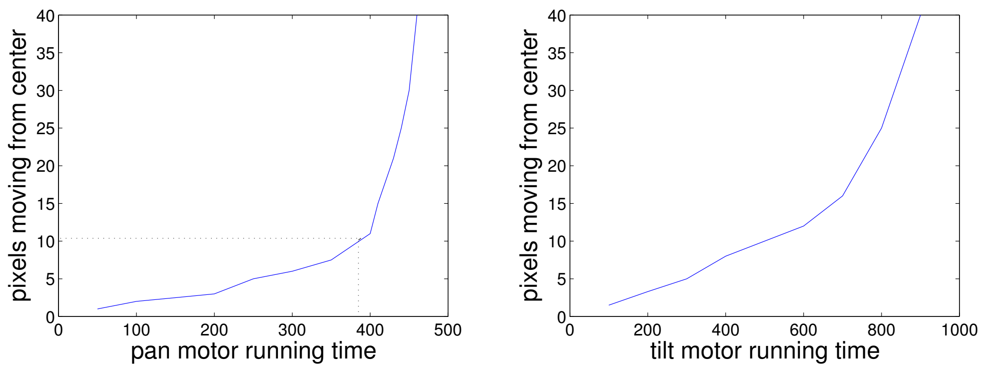

Figure 5. "running time-moving pixels" mapping table

3.4 PTU control method

It is difficult to control PTU both rapidly and stably at the same time. If we turn the pan-tilt unit faster, the unit may overshoot the expect position. But if we turn it slower, the unit will arrive at the expected position after too many steps.

In order to solve this problem, we lookup the "running time-moving pixels" mapping table: If we run the motor for certain time, the objects in the camera's center will move corresponding pixels to another position in the camera view, and vice versa. We record the motor running time and the moving pixels from center, then we can get the mapping table as **Fig. 5**.

According to this table ,we can lookup the time we should run motors after knowing the pixels depart from the center. For example, an eye is on the 10 pixels left of the W-Cam's reference point, so we will run the pan motor right for 380 ms. After that, the eye will appear at the reference point. If once is not enough, we can run motor in next cycle.

This control method is not very good in theory, but do well in practice. We expect better method more fast and stably in the future.

3.5 Auto zoom and auto focus

The imaging principle is described as below equation:

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \tag{3}$$

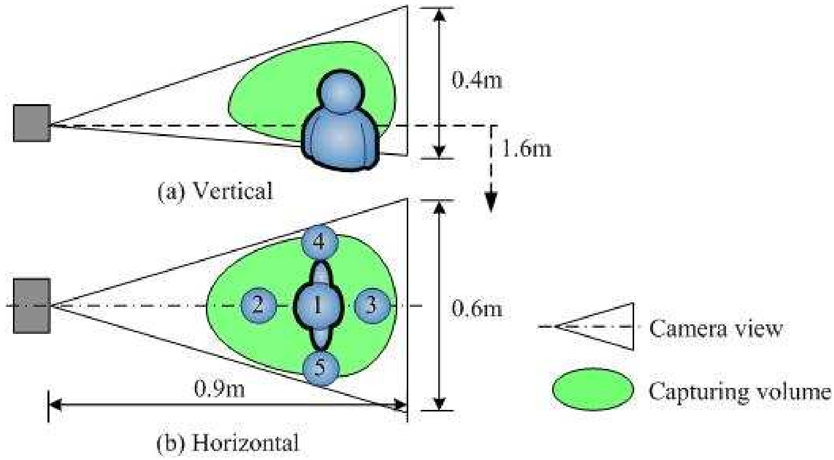

Figure 6. The capture volume and experiment design

Here,f is the camera focus, u is the object distance and v is the image distance. When the object distance u changes, we should change f or change v to keep the image not blurred. Both the two method can finish this function, but to change f can change the proportion of the object's image size at the same time, while, to change v seems much easier.

In our system, we use both auto-zoom and auto-focus to extend the capturing depth. Our camera has auto-zoom function driven by video frequency itself. Besides, we add auto-focus function: When people stand far, we set focus long and when people stand near, we set focus short, until the image is clear enough. Herein, the users' distance u can be estimated by the face rectangle size W or H with the same method with section 3.3, and the image quality is simply assessed by the high frequency weights. The auto-focus and auto-zoom function run at the same time and run parallel with the PTU moving, so this process takes little time.

3.6 image processing

After the clear iris image is acquired by N-Cam, the subsequent step is the image processing and iris recognition. Here, we use the existing algorithms, including the pulling and pushing method[14] and integro-differential operator[1] to precisely localize the iris, the local ordinal encoding algorithm to represent the iris feature[13] and hamming distance to do the feature matching with the registered feature templates[1]. These algorithms have been integrated in a dynamic link library (dll) by other people.

4. EXPERIMENT AND DISCUSSION

There are two important factors for evaluating the iris recognition system's performance, one is whether the images can be acquired easily and conveniently, and the other is the recognition rate, such as false matching rate (FMR) and false non-matching rate (FNMR). In this paper, we pay more attention to the image acquisition part and we have done some experiments on the capture volume, the tracking speed and also image quality which is related to the recognition rate.

4.1 Capturing volume

The capture depth is about 0.4 m to 0.9 m. If people move further, the iris diameter is less than 150 pixels, even if with the maximal focus. and if people move nearer, the iris image can not be focused.

The capture angle is mainly limited by the visual field of the wide angle camera, because the face must appear in the wide angle camera. The W-Cam's view is a pyramid of about 0.6 m width and 0.4m height in the distance of 0.9 m.

Summing up the capture depth and angle, the capture volume is actually the range showed in **Fig. 6**, which is an irregular shape, equivalent to the range of about $0.6\,\mathrm{m}\times0.4\,\mathrm{m}\times0.4\,\mathrm{m}$. If using longer lens and wider angle cameras, we believe we can get more capture range.

4.2 Tracking speed

We define the start of tracking when the W-Cam find user's face, and define the end of tracking when the N-Cam get the clear iris images. The iris image's sharpness is evaluated by the high frequency weights[1] and when the sharpness weights is greater than a set value, we think the image is clear enough for recognition. At the same time, the system sounds a beep and we record the time as the tracking time.

We fixed the device at about 1.6 m height and marked several positions and scale lines on the ground as **Fig. 6**, then let about 30 persons with different statures, different genders and different face sizes to perform the iris recognition. The test results are shown in **Table 1**.

Table 1. Average tracking time of different positions and different statures

Standing Position	**1**	**2**	**3**	**4**	**5**	Average time(s)
Height of 1.60 m	3	5	7	6	6	5.4
Height of 1.65 m	3	4	7	5	6	5.0
Height of 1.70 m	2	5	5	3	4	**3.8**
Height of 1.75 m	2	4	6	4	5	4.2
Height of 1.80 m	3	—	8	7	7	6.3
Height of 1.85 m	3	—	9	7	8	6.8
Average time(s)	**2.7**	4.5	7.0	5.3	6.0	**5.1**

From the **Table 1**, we can see that the average capture time (excluding the iris feature matching time) of all positions is about 5.1 seconds, but for the best positions, such as position 1, the average time is only 2.7 seconds, which is a satisfying result.

Persons of medium height will take less time, for example, the 1.7 m height persons take only 3.8 seconds average in all positions. Persons too tall or too low will take more time, for example, the 1.60 m and 1.85 m persons take more than 6 seconds, and even their faces can not be found in the W-Cam at the position 2. But we should remember, persons of medium stature is the majority, so the real average time must be less than the average time in the Table 1.

During experiments, we also find that eyes can not be detected in some frames sometimes, because the face detection and eye detection algorithms can not work well, when people's faces are too slanting or eyes are not widely open.

In conclusion, the tracking speed is satisfying when persons stand in the best position, and we can expect the average time to 2.7 seconds. Even in a little bad position, the average time is still 5.1 seconds. We believe this system can be applied in the future practice.

4.3 Image quality and recognition rate

In our system, we used the auto-zoomed camera, so most captured images are at focus and have good quality. **Figure 7** is an example picture we acquired from 0.9 m away with 75 mm focus. The image presents a clear texture, sufficient enough for identification tasks.

To explain the image quality quantitatively, we test the system's recognition rate in a low-scale population (about 30 persons). The system continuously captures the iris image, and generates the iris feature code and matches them with the registered feature templates. (The recognition algorithms are described at section 3.6.) Within 3 seconds, if the person is recognized right, the recognition succeeds, otherwise, we think the recognition fails.

About 30 persons with no glasses took part in our experiment, the false matching rate (FMR) is 0 and the false non-matching rate (FNMR) is 0.13 (only 4 persons can not be recognized with 3 second). But if we set the recognition time greater (such as 5 seconds), FNMR is reduced to 0.06 (only 2 persons can not be recognized).

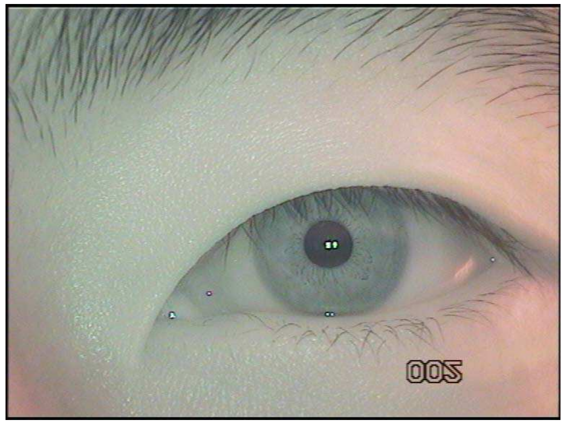

Figure 7. the quality of iris image

The main reason of the false non-matching is that the algorithms are not adapt to our system well and sometimes the registered image is not very representative.

The experiment result is not very good, but acceptable, considering that we didn't regulate the algorithm parameters at all. In the future we will regulate the recognition algorithms specially for our device and test the performance in a larger population again. The recognition rate must be much better.

Actually, many other factors will affect the system's performance, including illumination, lens, CCD quality and so on. In this paper, we don't want to discuss further, but only want to explain from an engineering view that our system can easily acquire iris images very well.

4.4 The advantage of low resolution cameras

The following section accounts for the advantages of the low resolution cameras for capturing iris image.

First of all, the low resolution cameras have high frame rate, generally more than 30 frame per second (fps). Contrarily, most high-resolution cameras have low frame rate for large data to transmit and limited bandwidth, for example, the camera Pulnex 4100 used by the "Iris on the move" system[10] has only a 15 fps rate, and the camera Mitsubishi used[8] only has less than 3 fps. A low frame-rate is a bottle-neck for iris images acquisition, since there might have few occasions in which the eyes region pass by the focus point. On the contrary, it is more likely that the eyes pass by the focus point with a higher frame-rate.

Secondly, on an financial aspect, the high resolution and high performance cameras are more expensive. For example, a camera with more than 4 mega pixels and more than 10 fps is nearly 10 thousand dollars, 30 times than the common 0.5 mega pixels cameras. It is an important factor to bear in mind.

Although the high resolution cameras cover a larger acquisition range, we consider it not as a critical issue since we have resolved it with the PTU. Anyway, our attempt on the non-cooperative iris image acquisition base on the low-resolution cameras is worthwhile.

5. CONCLUSIONS

In this paper, we introduced an active self-adaptive system for iris image acquisition. The system can automatically acquire iris image in large capture range and in a short time, but using only low-resolution cameras, processing image more fast and largely lowering hardware cost. The system integrated techniques of imaging schemes, face and eye detection algorithms, auto-focus and auto-zoom functions, and the PTU control methods. A prototype device has been established base on them.

We believe this work is a good attempt to make iris recognition more convenient and inexpensive.

ACKNOWLEDGMENTS

This work is funded by research grants from the National Basic Research Program (Grant No. 2004CB318110), Natural Science Foundation of China (Grant No. 60335010, 60121302, 60275003, 60332010, 69825105, 60605008), Hi-Tech Research and Development Program of China (Grant No.2006AA01Z193).

REFERENCES

1. John Daugman: How iris recognition works, IEEE Trans. Circuits Syst. Video Technol., Col. 14, no. 1, 21-30, 2004
2. Tony Mansfield, Gavin Kelly, etc.: Biometric Product Testing Final Report. Technical Report, Centre for mathematics and Scientific Computing, National Physics Laboratory, Teddington, Middlesex, UK, 2001
3. James L. Cambier, John E. Siedlarz, inventors; Iridian Technologies Inc.: Portable authentication device and method using iris patterns , US patent 6,532,298, 2003
4. Takahiro Oda, Yuji Ohta, inventors; Oki Electric Industry, assignee: Photographing device iris input device and iris image input method. US patent 6,850,631. 2005
5. Tomoyoshi Nakaigawa, Ken Ikoma, inventors; Matsushita Electric Industrial Co., assignee: Iris image pickup apparatus and iris authentication apparatus. US patent 7,068,820. 2006
6. Woo Hee Lee, Aekyung Yang, Jang jin Chae, inventors; LG Electronics, assignee: Apparatus and method for adjusting focus position in iris recognition system US patent 7,095,901. 2006
7. Keith Hanna, etc., inventor; Sarnoff Corporation, assignee: Fully automated iris recognition system utilizing wide and narrow field of view. US patent 6,714,665. 2004
8. Guodong Guo, Michael J.Jones: A System for Automatic Iris Capturing. Mitsubishi Electric Research Laboratories, 2005, http://www.merl.com/publications/TR2005-044
9. Dominick LoIacono, James R.Matey, inventors; Sarnoff Co., assignee: Method and Apparatus for obtaining iris biometric information from a moving subject. US patent 20060274919. 2006
10. James R. Matey, Keith Hanna, etc.: Iris on the move: Acquisition of Images for Iris Recognition in Less Constrained Environments, Proceedings of the IEEE, Col. 94, No.11 :1936-1947, 2006
11. Paul Villa, Michael Jones: Rapid object detection using a boosted cascade of simple features. IEEE Conference on Computer Vision and Pattern Recognition, Col 1:511-518, 2001
12. Craig Fancourt, Luca Bogoni, ect.: Iris Recognition at a distance. AVBPA 2005, Vol. 3546: 1-13. Springer-Verlag, Berlin Heidelberg, 2005
13. Zhenan Sun, Tieniu Tan, ect.: Robust Encoding of Local ordinal measure: a general framework of iris recognition, Proc. of Intl. Workshop on Biometric Authentication(BioAW),Vol.3087,270-282,2004
14. Zhaofeng He, Zennan Sun, Tieniu Tan: Iris localization via pulling and pushing. IEEE Proc. of Intl. Conf. on Pattern Recognitino(ICPR'06),1-4,2006
15. Sowon Yoon, Ho Gi Jung, ect.: Non-intrusive iris image capturing system using light stripe projection and pan-tilt-zoom camera. IEEE Proc. Computer Vision and Pattern Recognition(CVPR'07),1-7,2007