Abstract

This paper proposes non-intrusive iris image capturing system, which consists of pan-tilt-zoom camera and light stripe projection. Light stripe projection provides the position of user. After panning according to user’s position, AdaBoost-based face detection finds tilt angle. With user’s position and tilt angle, zoom and focus position are initialized. User’s position replaces 2D face search with 1D face search. Exact zoom and focus position enable fast control and narrow search range. Consequently, experimental results show that proposed system can capture iris image within acceptable time.

1. Introduction

The best biometric characteristic has five qualities: robustness, distinctiveness, availability, accessibility and acceptability. Robustness means unchanging on an individual over time. Distinctiveness means showing great differences among each individual. Availability means the entire population has this measure in multiples. Accessibility means this measure can be easily imaged to an imaging sensor. Finally, acceptability means perceived as non-intrusive by the user [1, 2].

Iris recognition satisfies three qualities: robustness, distinctiveness, and availability. Iris patterns are fixed from about one year of age through life [3]. According to [4], iris pattern has 244 independent degrees of freedom. This means that the probability of two different irises agreeing by chance in more than 70% of their phase sequence is about one in 7 billion. And every person has iris patterns and two irises have distinctive patterns.

However, iris as a biometric feature has not been used as widely as fingerprint and face. In view of accessibility, obtaining in-focus iris image is not easy. For iris recognition system using fixed focus lens, the system has to wait until users set their position in well-focusing range.

For iris recognition system using auto-focus imaging sensor, the system can capture iris images more actively, but it still has small operating range so that the iris images can be acquired when a user closely approaches to the system. Both systems result in low throughput. In view of acceptability, a user has to stay very still in front of the camera and look at the camera without any motion and blinking. This causes user inconvenience.

To overcome the limitation of accessibility and acceptability of conventional iris recognition systems, new types of iris recognition system are suggested: Sarnoff Corporation [4], Mitsubishi Corporation [5], and authors [6].

Sarnoff Corporation suggests an iris recognition system for the users with normal walking pace through a portal [4]. Their system continuously captures the images of expected region with high-resolution while a user passes through the portal. Average 1.5 in-focus iris images can be captured. Their system has high throughput since iris images of 20 persons per minute can be obtained ideally. However, their system still has constraints on user’s position and pose. Users with great height or small height can not be accepted in this system. Although they suggest modular component for covering the height variation, the cost may increase proportional to the operating range they want to cover. This may cause some difficulties in use for children and the disabled. Also, users have to follow some instructions such as look straight ahead and open the eyes widely. That is, Sarnoff Corporation’s system increases accessibility, but still has low acceptability.

Mitsubishi Corporation suggests an iris recognition system combined high resolution camera with wide-field-of-view camera using a pan-tilt unit [5]. Wide-field-of-view camera scans the operating range for detecting a user’s face, and then estimates the distance from the face to the camera. This distance helps high resolution still camera to change direction and snap a photo. Their system provides wide operating range determined by field-of-view of wide angle camera and pan-tilt range. But scanning based on face detection takes time and distance.
estimation based on facial feature extraction is not accurate because each person has different size and interval of facial features.

Authors suggest a pan-tilt-zoom based iris recognition system [6]. Panning and tilting can cope with width and height variation, respectively. By zooming, the distance variation between camera and user can be extended greatly. This system provides wide operating range and allows pose variation of user so that it increases acceptability. However, its throughput is low because scanning whole operating range to detect user’s face takes a lot of time and zoom factor itself based on face detection does not give focus lens position to obtain in-focus iris image. That is, initial position of focus lens is difficult to determine by only zoom factor without distance between camera and user. As a result, focus lens moves almost full range to search optimal position. And if a user moves after fully zoomed and the face is out of view, then the system should to go to the initial step and repeat whole process to know the position of user even though the location of face is close to previous position.

This paper maintains user acceptability of [6] using pan-tilt-zoom camera, improves throughput and copes with position variation by adapting wide-field-of-view camera and light stripe projection which can detect the position of user all the time. Position estimation by light stripe projection has high accuracy. Therefore scanning whole area is not needed. Position of zoom and focus lens can be initialized by the distance and the search range of focus lens position can be greatly reduced. Also since this system always knows the position of user, it can catch user’s face despite user’s movement. As a result, this system achieves accessibility as well as acceptability.

2. System Configuration

2.1. Pan-Tilt-Zoom Guided by Light Stripe Projection

This system provides unconstrained user environments in position and pose. By panning and tiling, width and height of operating range are 1m and 1m, respectively. By zooming, depth of operating range is 1m. This means that the system can capture the iris image of the user in this volume for recognition. To obtain the iris images robustly, high resolution image sensor is used. Thus an image of full face after zooming is enough to acquire iris image with medium quality resolution for recognition [7]. This guarantees at least one iris to remain in an image even though the user moves their body.

For process efficiency, this system adapts light stripe projection with wide-field-of-view camera. Fig. 1 shows proposed system configuration. Fig. 3 shows the pan-tilt-zoom camera coordinate system. The coordinates of wide-field-of-view camera are represented by X, Y and Z. The coordinates of pan-tilt-zoom camera are represented by X’, Y’, and Z’. Using light stripe projection, X-directional and Z-directional positions of user in operating range are detected. X-directional position determines pan angle. Z-directional position is used for calculating of the distance D between pan-tilt-zoom camera and user’s face and corresponding zoom and focus lens position. It reduces scanning area from 2-dimensional area, i.e. 1m × 1m (width × height), to 1-dimensional area, i.e. 1m (height). The search range of focus lens position is also reduced thanks to calculated initial position. The entire concept is depicted in Fig. 2.
2.2. Flow Chart

Fig. 4 shows overall procedure. Line detection from wide-field-of-view camera and calculation of $X$ and $Z$ position of user are in Section 3. Panning by $X$ position and tilting with face detector are required to center the face in an image. Then the distance $D$ between pan-tilt-zoom camera and user is calculated. Based on $D$, proper zoom and focus lens position are exactly calculated from the relationships of distance-zoom position and zoom position-focus position which is estimated from empirically measured data. Given small search range of focus lens position, the best in-focus iris image is acquired. This is dealt in Section 4.

3. Estimation of User Position

In this section, we explain how to detect light stripe on the user’s leg and get $X$ and $Z$ position of the user from detected light stripe.

3.1. Light Stripe Detection

Using an optical narrow band pass filter, light stripe generated by light projector can be easily extracted in unconstrained environment. The remaining problem is to distinguish the light stripe on the user’s leg. In this paper, CC (Connected Component)-based background subtraction is used. This is robust to small change of background or camera position than pixel-based background subtraction which can result in strong disturbance when the light stripe is slightly moved.

Input image is converted into binary image by adaptive thresholding. A light stripe on user’s leg sometimes consists of several CC. To make these line segments as a line, dilation is performed with $6 \times 2$ kernel which connects horizontal directional segments. Finally, CCs with small number of pixels are removed since they are considered as noise.

CC-based background subtraction eliminates background by subtracting background CC image from input CC image. Background CC image is stored before the start of operation. The only CCs without any overlap with background CCs are remained. It is expected to be the stripe on the leg. In Fig. 5, (c) shows the input CC image and (d) shows the detected CC on the user’s leg.

3.2. XZ Estimation

By the principle of light stripe projection, 3D information of the light stripe can be calculated [8]. The meaningful information is $X$ and $Z$ position of the detected light stripe. We are able to calculate $XZ$ position of user.
from the center of detected light stripe (Fig. 5(e)).
To calculate $XZ$ position, the first step is the calibration of wide-field-of-view camera. A fisheye lens is used here to observe wide operating area. Since wide angle lenses have radial distortion, the undistorted image is obtained by refined radial distortion model [9]. In Fig. 5, (a) shows input image without optical filter, which is included only for illustration, and (b) shows the undistorted image. Only the center point of light stripe is transformed to undistorted image coordinates instead of whole image rectification.

The second step is the estimation of parameters: $\alpha$ and $b$ in Fig. 3. Assume $\rho = 0$ which implies the light stripe is parallel with X axis. Then, original formulas for $X$, $Y$ and $Z$ in [8] reduce to (1), (2), and (3) as below:

$$X = \frac{x \cdot b \cdot \tan \alpha}{f - y \cdot \tan \alpha}$$
$$Y = \frac{y \cdot b \cdot \tan \alpha}{f - y \cdot \tan \alpha}$$
$$Z = \frac{f \cdot b \cdot \tan \alpha}{f - y \cdot \tan \alpha}$$

Given $f$ and many pairs of a point in undistorted image $(x, y)$ and $Z$, we estimated $\alpha$ and $b$ by least-square estimation.

Using $\alpha$ and $b$, a new point $(x, y)$ can be transformed to $(X, Y, Z)$. Since $X$, $Y$, and $Z$ are wide angle camera coordinate, transformation to pan-tilt-zoom camera coordinates is required as shown in (4). $\gamma$ is tilting angle of wide angle camera about $X$ axis, $h_{ptz}$ is height of pan-tilt-zoom camera and $dZ_{\text{wide-\text{ptz}}}$ is $Z$-directional offset between origins of two camera.

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma & -\sin \gamma \\ 0 & \sin \gamma & \cos \gamma \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} 0 \\ h_{ptz} \\ dZ_{\text{wide-\text{ptz}}} \end{bmatrix}$$

Finally, the $X'$ and $Z'$ position of user in pan-tilt-zoom camera coordinate are obtained. $Y'$ information is not necessary in this case because this indicates the height of light stripe on the leg. High accuracy of $Z'$ position is revealed by the fact that error range of $Z'$ is 1–15mm.

4. Pan-Tilt-Zoom Control

In this section, we deal with panning based on $X'$ and $Z'$, tilting based on face detection, and zooming and focusing by estimating the distance between pan-tilt-zoom camera and user’s face.

4.1. Pan Angle Determination

Pan angle is directly calculated by $X'$ and $Z'$ as shown in (5) and Fig. 6:

$$\theta_{\text{pan}} = \tan^{-1} \left( \frac{X'}{Z'} \right)$$

Since pan angle is deterministic, scanning whole operating range in all directions is not required. Instead, $Y'$-directional search at the fixed pan angle is needed.

4.2. Tilt Angle Measured by Face Detection

After panning, upward scanning to find face is performed. In Fig. 7, (a) shows $Y'$-directional scanning direction and (b) shows the images at some tilt angle. ‘Stop’ command of $Y'$-directional scanning is occurred when the face is detected. The top image of Fig. 7(b) shows result of face detection.

Face detection based on AdaBoost algorithm is used here since it is fast and accurate [10]. For this scanning pan-tilt-zoom camera moves upward direction and the tilting angle is at most 33.69° for covering 1m height variation at the closest distance from camera, i.e. 1.5m.

From the position of tilt when face is detected and centered in an image, actual distance $D$ from pan-tilt-zoom camera to user’s face is obtained by tilt angle $\theta_{\text{tilt}}$ as shown in (6) and Fig. 6.

$$D = \frac{Z_d}{\cos \theta_{\text{tilt}}}$$

Figure 6: Pan angle $\theta_{\text{pan}}$ and $D$. Pan angle $\theta_{\text{pan}}$ by $X'$ and $Z'$, and $D$ by $\theta_{\text{tilt}}$. 
4.3. Zooming and Focusing

Estimation of $D$ sets initial position of zoom and focus lens. Since iris size of each person is almost same and the required resolution of iris image is known [7], the magnification $M$ is calculated. Given $D$ and $M$, optimal position of zoom and focus lens position can be specifically determined.

The relationship among distance $D$, zoom position $Zoom$ and focus position $Focus$ is obtained during calibration procedure. While we changed $D$ by 5cm, we found a zoom factor such that face image have the same magnification. With the zoom factor, focus lens position was changed by 10 steps. The optimal focus lens position is determined by finding the position with maximum focus measure value, which is evaluated by the kernel in [11].

Fig. 8(a) shows the relationship between $D$ and zoom position, and Fig. 8(b) shows the relationship between zoom and focus position. Under constant magnification, zoom position and $D$ has inverse relation. To estimate feature of our lens, we set the model as (7). The model parameter $z_1$ and $z_2$ were estimated by SVD (Singular Value Decomposition) as shown in Fig. 8(c).

$$\text{Zoom} = z_1 - \frac{z_2}{D}$$  \hspace{1cm} (7)

The relation between zoom and focus position can be modeled linearly like (8). Using least-square, we estimated $f_1$ and $f_2$ as shown in Fig. 8(d).

$$\text{Focus} = f_1 \cdot \text{Zoom} + f_2$$  \hspace{1cm} (8)

The most efficient way to focus on the object is to know the distance from camera to object and zoom factor. Between two features, the distance is more important than zoom factor. According to fundamental theorem, the probability density function (pdf) of output which passes though a function $g(x)$ is inverse proportional to magnitude of the derivative of $g(x)$ [12]. If we can know the distance with a certain error bound, estimated zoom factor usually has smaller error bound (Fig. 9(a)). On the other hand, if we know the zoom factor with a certain error bound, estimated distance has large error bound (Fig. 9(b)). Therefore, the accurate distance makes possible faster operation by reducing the search range of focus lens position than the focusing mechanism based only on zoom factor.
5. Depth of Field, Depth of Focus and Operating range

Practical iris recognition range is defined in [6]. Only when Hamming distance is smaller than threshold, the iris can be authenticated correctly. At every 5cm from pan-tilt-zoom camera, position of focus lens is changed within the interval including optimal focusing position (Fig. 10(a)–(e)). Then depth of focus, depth of field, and operating range is obtained. Fig. 10(f) is a measured Hamming distance example with respect to focus lens position. As a result, depth of focus is 500–2000 steps, which implies that error of focus lens position within this range is acceptable. A system with wide depth of focus does not require precise focusing algorithm. And depth of field is 5–9.5cm so that this system is robust to small movement of user. Fig. 11 shows results of measured depth of focus, depth of field and operating range.

6. Experimental Results

6.1. Capturing Time

Two main factors reduce the capturing time: pan angle calculated by $(X', Z')$ and pre-deterministic zoom/focus position. Instead of scanning the whole operating range in order to find user's face, upward scanning just once in $Y'$ direction is performed. Time for tilting and panning depends on the tilt and pan angle, respectively. However, experimental results show that time for tilting+panning does not exceed 2 seconds. Time for zooming and focusing is almost constant. They are deterministic because zoom and focus position are calculated by calibrated model without scanning. Fig. 12 shows the nine test positions covering the whole operating range. Right hand side of Fig. 12 lists the measured panning, tilting, and zooming+focusing time for each test position. Total capturing time is around 4 seconds.
6.2. Accuracy of Focus Lens Position

Estimated distance $D$ between pan-tilt-zoom camera and user’s face initializes zoom and focus position. Experimental results confirmed that zoom position was properly selected such that face had almost the same size. By examining the focus measure corresponding to the initial focus position, we evaluated how well the initial position was properly selected. Fig. 13 shows the measured focus value with respect to focus lens position and the focus measure corresponding to initial focus position. It can be observed that the initial focus position has very high focus measure value.

![Figure 13](image_url)

Figure 13: Focus measures according to focus lens position and given initial focus lens position. (a) overall focus measurements varying focus lens position (solid line) and initial focus lens position given by $D$ (star), and (b) enlarged graph of the dashed rectangle in (a).

7. Conclusions

In this paper, we propose a novel iris capturing system based on pan-tilt-zoom camera and light stripe projection. This system provides user convenient environment as well as acceptable processing time about 4 seconds. By $X'$ and $Z'$ information, pan angle is determined. After tilting for face detection, the distance $D$ between pan-tilt-zoom camera and user’s face can be calculated. Because $D$ can determines the zooming and focusing position deterministically, zooming+focusing time can be drastically reduced. Furthermore, initial zoom and focus position are expected to guarantee high quality iris image.

Acknowledgments. This work was supported by the Korea Science and Engineering Foundation (KOSEF) through the Biometrics Engineering Research Center (BERC) at Yonsei University.

References