Master Thesis (Double Degree Program)

Eye Region Detection by Likelihood Combination for Improving Iris Authentication

Entrance Year: 2016

Thesis Submission: August 17, 2017

Graduate School of Engineering

University of Miyazaki

Thae Su Tun

Supervisor: Professor Masayuki MUKUNOKI

CHAPTER 1. INTRODUCTION

As the public safety becomes the primary concern, person identification is considered as an important issue to achieve a secure and safe society. A variety of systems require reliable personal identification schemes to either confirm or determine the identity of an individual requesting their services.

Approaches for person identification can be roughly classified into two types as active and passive approaches [1]. In the active approach, special equipment needs to be installed to transmit identifying information of an object. RFID (Radio Frequency Identification) tags are examples of the active approach. The merit of the active approach is that it automatically identifies the observed objects. A demerit of such approach is the cost due to the use of battery. Moreover, since it can be easily tapped or intercepted, there is the potential risk for unauthorized access to the data on the RFID tag when transmitting the data between the reader and the tag. Furthermore, RFID tags can be clone easily with available equipment. The passive approach extracts identification information using environmentally-fixed sensors such as video cameras and microphones. The merit of the passive approach, such as iris and fingerprint authentication, is that it is less possible for unauthorized access as it cannot be easily intercepted by others.

In the passive approach, there are three major methods for authentication purpose [2]. The two conventional methods are based on the person's possession of token, such as key or ID card, and possession of knowledge, such as password. These two methods are not secure enough because the key or ID card can be stolen and/or imitated by others. In the password authentication, the security is totally based on the strength and confidentiality of the password. The third method for authentication is based on person' s biometric traits. Biometric recognition becomes a strong link between a person and his/her identity because biometric traits cannot be easily shared, lost or duplicated. There are two types of biometric features, physiological features and behavioural features. Examples of physiological features are finger print, palm print, face and iris. Examples of behavioural features are hand writing, key stroke and gait.

Iris recognition is considered to be one of the most accurate methods in all biometrics authentication methods as iris has unique features and even the irises of identical twins are not the same. Furthermore, iris recognition is a safe, reliable and convenient technology for personal recognition. It is reliable as iris is an inner organ in our eyes and protected by the eyelid and eyelash. It is rarely hurt and the error caused by the scar will

rarely occurs as in the case of fingerprint, palm and face. Moreover, our irises are stable from one-year-old until the end of life. So, iris is more stable and robust compared with other biometrics as fingerprint, palm and face.

Iris recognition is widely used in many application areas such as secure financial transactions, national border control, driving licenses, credit card authentication, internet security, computer login, room access control, ticketless travel, cell phone and other wireless-device-based authentication. As it doesn't need too much difficult for data acquisition, iris recognition is easy to implement than other biometrics.

A general iris recognition system is composed of five main steps, image acquisition, iris segmentation, iris image normalization, feature encoding and matching. The performance of an iris recognition system mainly depends on the quality of captured iris images. In data acquisition step, eye images can be captured under constrained or less constrained environment. The iris images under constrained environment are captured at a fixed short distance and fixed arranged lighting condition. The users are constrained to stay at a fixed location of the head and starring at a spot. The resulting images taken under such environment have good quality. Iris images captured within a specified shorter distance are clear enough. On the other hand, few constraints are imposed on the users under less constrained environment. Since the freedom of user's movement and position increase, the variation of images taken under such situation also increases. These changes lead to differences in resolution, illumination and eye poses between different acquisition. When the acquisition environment varies, the similarity between irises of the same person eye decreases. Furthermore, the iris in the image appears small compared with ones taken at a short distance. Iris segmentation tends to find the best eye candidate and segment the iris region in an image. For images where iris appears as a small part and suffers from some bad lighting condition and illumination, the segmentation procedure segments wrong organs as an iris. The possible way to avoid wrong segmentation on other organs such as glasses, nose, hair and nostrils is to detect the eye region before proceeding other steps as segmentation, normalization, encoding and matching.

In this research, eye region detection is firstly performed by our proposed method for face images taken at a distance before the iris segmentation stage. The cut eye region images are given as input to an iris authentication system. If the eye region is detected firstly and giving this detected eye region as the input, it doesn't need to search the eye region all over the face image for segmentation and it is less possibility to wrongly segment other organs or glasses as the iris.

We propose a new eye region detection method combining two likelihood images. The first likelihood image is created by AKAZE [3] feature matching and the second by template matching. These two likelihood images are combined and the maximum pixel value location is searched from this combined likelihood image. The eye region is cut based on this position.

We perform the analysis of the proposed eye detection method using the public data set(CASIA-v4-Distance) [4]. The experimental results show that detection of eye region before segmentation can improve iris authentication performance.

This thesis consists of six chapters. This chapter provides an introduction of the research with respect to the objective of the research. Chapter 2 describes the background of a general iris recognition system. Chapter 3 includes a literature review on some related works and problems in the previous works. Chapter 4 gives the detail explanation of our proposed method and Chapter 5 presents the experimental results and the discussion of our research. Chapter 6 concludes this thesis by considering the insight obtained from our research and by recommending some future works.

CHAPTER 2. APPLICATION BACKGROUND

2.1 Biometric Authentication Systems

Person identification is the process of matching a physical person with the information that belong to this person. The purpose is to determine the pre-established individual identity or verify individual identity claim with the aim of security. Person identification methods are changing from conventional methods to biometrics identification methods for security purpose. Person identification using biometric features has more advantages than conventional methods like key, ID card and password.

The term biometrics is a computerized system implemented using biometric features for authentication. A biometric system provides automatic recognition of an individual based on unique features or characteristics possessed by the individual. Biometric recognition has been popular as a reliable identity management system since only the user oneself is necessary for authentication.

There are two major types of biometric features, the first one is behavioural and the second physiological features. Behavioural features are based on the behavioural pattern of human that is the general physical traits such as dynamic keystroke, gait, hand writing/ signature and voice. The behavioural features are not robust enough as they may be changed over time. The physiological features are based on relatively stable characteristics and shape of the human body which are difficult to alter and imitate by others. The DNA (Deoxyribonucleic acid), teeth, hand geometry, fingerprint, palm print, face, retina, iris are some examples of physiological features. The physiological features are more robust than the behavioural features as the physiological features are parts of human body and almost unique for all persons.

Since biometric traits have highly reliable and unique features, biometric authentication systems are widely used in a wide application domain in industrial fields. Each biometric trait has its own strength and weakness. The effectiveness of biometric is measured based on the recognition performance accuracy which need to meet, user acceptability, stability(robustness), resource requirements (easiness for data acquisition) and cost in all the applications.

Table 1. Comparison of some biometric traits

	Biometric Traits	Performance	Permanence	Collectability	Cost
Behavioural features	Gait	Low	Low	High	High
	Signature/ Hand writing	Low	Low	High	Low
Physiological features	Fingerprint	Medium	Medium	High	Medium
	Palm print	Medium	Medium	High	High
	Face	Low	Medium	High	High
	Retina	High	High	Low	High
	Iris	High	High	High	Medium

Each biometrics have its own benefits and appropriateness for different application domains. Fingerprint, teeth and DNA are more useful for criminal forensic cases. For accurate person identification, iris and retina based authentication are more reliable because of its high performance. There are three kinds of biometrics such as vision based, audio based and the last one is the use of human body parts. In the vision based biometrics, the images or video taken by camera or sensor are used as input for identification. signature/ handwriting, hand geometry, fingerprint, palm print, face, retina and iris are vision based biometrics. Since a microphone or other audio sensor is needed to record a person's voice, it is an audio based biometric. DNA and teeth are biometrics that human body parts are needed for person identification. The Table 1 briefly describes the comparison of some vision based biometric traits' strength and weakness in terms of performance, permanence, collectability and cost. We will discuss the advantages and disadvantages of each biometric feature based on Table 1 in the following.

Person identification based on gait of human has advantages of non-invasive biometric and can recognize a person at a distance. But person wearing a rain coat or trench coat and changing the foot wears, walking surfaces uphill, downhill can cause the problem for person identification. Moreover, it may change due to tiredness, health and age. Since

gait analysers are generally huge and need a well-equipped room to analyse an individual, it is an expensive biometrics.

In signature recognition, it is non-invasive, cheap and needs little time for verification. But, the users need to sign in a consistent manner for accurate verification. It has higher error rate than other biometric traits as there is large intra class variation. Moreover, the physical and emotional effects of users during signing may cause large variation and affect the recognition performance. Person identification using hand writing is also easy to use and non-invasive for data acquisition but it is not accurate and it can be imitated by others.

Faulds, Herschel and Galton [5] discovered the distinctive nature of fingerprints in the late 19th century. At that time, almost all law enforcement agencies in the world used fingerprints for criminal and forensic identification. Human fingerprints are detailed, nearly unique, difficult to alter for an individual. The recovery of fingerprint in a criminal scene is an important method in forensic science. The fingerprint technology is cheap, easy to use and user friendly. However, dryness or dirty of the finger's skin can make errors at the time of enrolment. Furthermore, as fingerprints are outside organs and it can be harmful and damaged by other injuries, it is not too much reliable for authentication.

Palm print recognition is a biometric technology based on the patterns of palm of human hands. Since palm print has much larger area than fingerprint, more distinctive features can be captured and it is suitable for person identification than fingerprint. The palm print scanners are large and expensive as they need a larger area to capture than fingerprint scanner.

Li and Jain [6] proposed a person identification method based on face. Face recognition technology has many benefits such as its convenience and social acceptability and it is easy to use as it can be performed without user knowing. One of the key disadvantage of face recognition is that it cannot correctly identify for identical twins. Furthermore, a person' face may be changed by surgical operation and along age. The appearance and environment changes such as shaving beard and wearing eye glasses can lead to false rejection for identification.

Biometrics using retina has very high accuracy as it cannot be replicated by others but it is very expensive. Users of this technology have complained that the enrolment process is uncomfortable and invasive. Users are afraid that the infrared light used in the retina scan damages their vision [7].

A new system for person identification based on iris patterns was proposed by Daugman [8]. Iris recognition has high accuracy and data acquisition is not so much difficult as retina. Iris recognition has many goodness compared with other biometric techniques as it is unique and rarely harmed by injuries. Iris recognition is reliable and robust, so it is conventionally used in many real world applications.

2.2 Iris Recognition System

Iris recognition is a kind of biometric technologies based on the physiological characteristic, iris. Iris is an externally visible internal organ protected by eyelid, eyelash and cornea. The human iris is the annular part between pupil and sclera. Iris has texture features such as contractile furrows, crypt, collarette, stroma fibres, wolfflin nodules and nevi. Iris textures are distinct for each person; thus iris is unique for individuals. An example of human iris image is shown in Figure 1 downloaded from Dr. Jan Drews' work [9]. The probability of two persons' irises being the same is lower than 10^{-35} [10]. Even in identical twins, their irises are totally different.

Iris is robust and errors caused by scars will seldom happen like fingerprint and palm. Moreover, our irises are mature when we were one-year-old and would not change in our life. Users don't need to touch the devices or sensors and it is comfortable for data acquisition. Iris becomes a living password which cannot be lost, forgotten and copied by others.

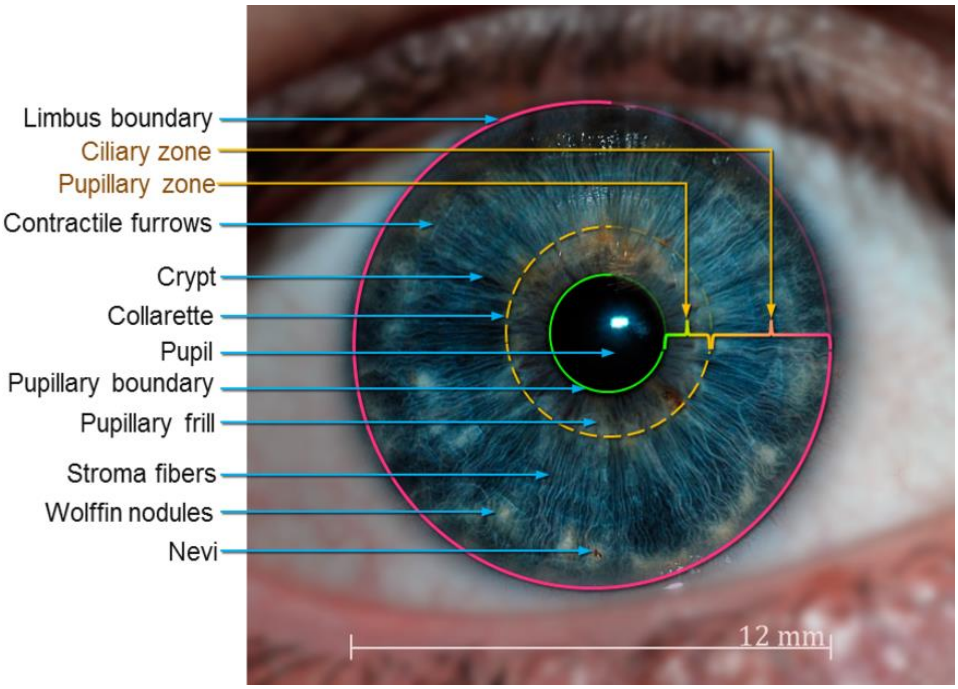


Figure 1. An example of human Iris image [9]

Since iris has more advantages as uniqueness, robustness and high performance, it is extensively used in numerous applications like national border control, securimetrics (portable handheld cameras for military and police), anti-terrorism (e.g. screening suspects at airport). Iris recognition technology is employed in several countries for controlling border access such as Amsterdam Schiphol airport in Netherlands, twenty-nine national airports in Canada, and also in the United States. This method has been deployed in many hospitals to protect steeling the new born infants from thieves. Iris recognition are also used in patient identification in order to accurately identify a patient upon entry into a healthcare facility and for their utility to quickly and accurately identify patients' current location in the hospital.

The first automatic system to identify people based on iris texture was developed by John Daugman [8] in the last decade of 20^{th} century. In the Daugman's iris recognition system, he divided the system into the four main stages: iris segmentation, iris image normalization, feature encoding and matching. Since 1993, Daugman has proposed several techniques that can improve the performance given by his first iris recognition algorithm.

The iris recognition system proposed by Daugman was followed by Wildes' system in 1997 [11]. The Wildes' work included new proposals for iris acquisition, segmentation, normalization, feature extraction and template matching. Although Wildes' segmentation is considered to be more stable to noise, Wildes' approach leads to higher EER (equal error rate) and more complex compare to Daugman's system.

Several researchers have proposed alternative solutions in order to improve the performance of the iris recognition system. In the research biometric community, free iris recognition systems are provided and each module can be studied and examined independently. The most widespread open source iris recognition system was supplied by Masek and Kovesi [12] in 2003. It was developed as a part of Masek's master thesis inspired from the first iris recognition system proposed by Daugman.

In recent years, NIST provided an open source iris recognition system in less constrained acquisition environment, namely VASIR (Video-based Automated System for Iris Recognition) [13]. VASIR is the first public open source iris recognition system for videos taken at a distance.

In 2007, an open source iris recognition system based on Daugman's approach, called OSIRIS (Open Source for IRIS) was proposed by Guillaume Sutra, Bernadette Dorizzi and Sonia Garcia-Salicetti but it is not available online [14]. The OSIRIS version 2 is the first version available online in 2009. But it is less used due to the weakness in the

segmentation stage. Then, the segmentation module is modified in the last OSIRIS version 4. In this research, we use OSIRIS version 4 as an example iris authentication system.

2.3 OSIRIS (A Reference system for IRIS authentication)

OSIRIS is an open source reference system for iris authentication system developed in the framework of the BioSecure project [15]. This version of OSIRIS version 4.1 was developed by Guillaume Sutra, Bernadette Dorizzi and Sonia Garcia-Salicetti in Telecom Sud Paris (previously GET-INT).

The following Figure 2 describes the overview of OSIRIS. OSIRIS is mainly encompassed with four processing modules: segmentation, normalization, encoding and matching. Firstly, an image is given as input to be used in segmentation and normalization. The segmentation stage produces the contour parameters and the segmented image. The normalization stage transforms the segmented iris region into normalized image and generates normalized mask to be used in matching stage. The normalized image is filtered by Gabor filters and iris template is generated to be used for matching stage. The Hamming distance between two iris codes are computed in the matching stage.

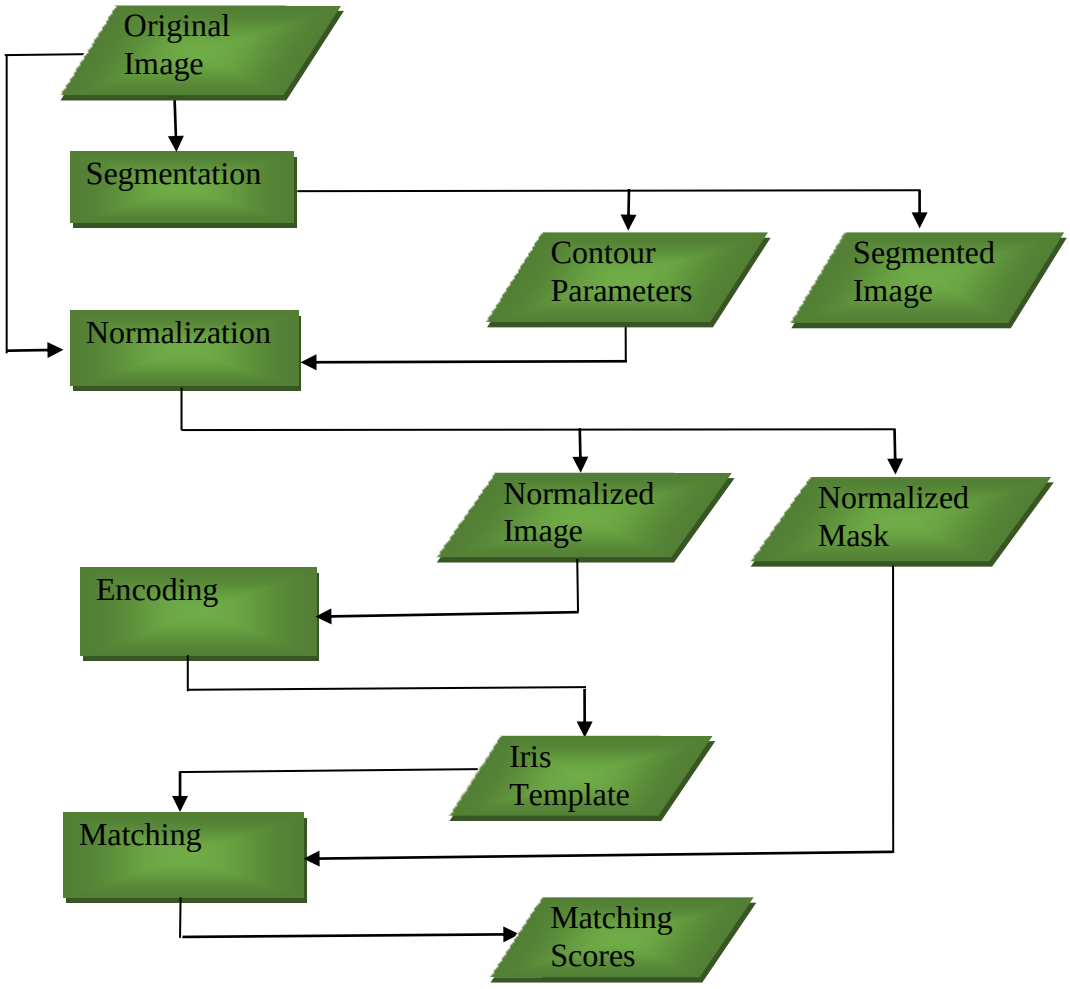


Figure 2. The diagram of OSIRIS reference system

2.3.1 Segmentation

Iris segmentation is the first key step of the whole iris recognition as the performance of iris recognition systems highly depends on segmentation. The objective is to separate the usable iris pattern from other parts of the eye such as eyelids, eyelashes and the noise such as illumination on the iris region. The segmentation step provides two contours (pupil and iris) that will be used by the normalization step. The segmentation stage aims to classify iris and non-iris pixels. The pupil and iris circles are searched by least-squares based circle fitting method applied to the coarse contour detected by the Viterbi algorithm. The segmentation step generates a binary mask of iris where on-pixels belong to the iris, and off-pixels do not belong to the iris which can be used in the matching stage. It is also an optional input for the normalization stage. But it needs to be normalized before matching stages.

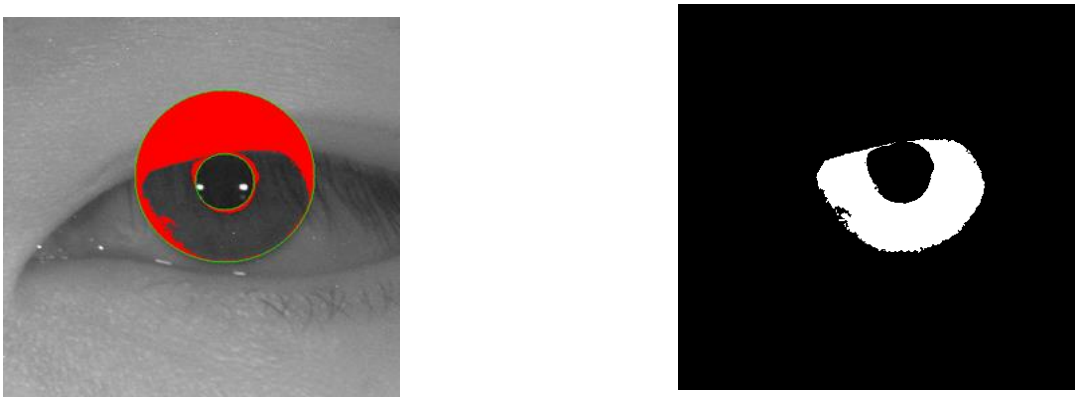
The segmentation stage aims at finding accurate contours of iris, that is inner boundary (pupil/iris) and outer boundary (iris/sclera). The pupil is localized by two criteria. These two criteria are “pupil is the disk shape black area” and “iris has circular edge contours”.

The first criterion corresponds to finding the minimum value in the image filtered by a disk shaped kernel of size *r*-by-*r* by summing all the pixels in a disk-shaped mask. A normalization is applied in order to range the resulting values between 0 and 1, and the result is inverted so that 0 corresponds to a good pupil candidate, and 1 to a bad pupil candidate.

The second criterion is achieved by four main steps. (1) Horizontal and vertical gradient maps of the image are computed using the Sobel operator and (2) two ring-shaped gradient kernels of size *r*-by-*r* is built (3) Filtering each map by each kernel respectively and (4) find the location of the maximum values in the sum of both filtering. This second criterion leads to values between 0 (bad pupil candidate) and 1 (good pupil candidate). These two values are added and the position which has the maximum value is found as the pupil centre.

Two sets of parameters corresponding to the pupil and the iris coarse contours (the number of point in the pupil and iris coarse contours, and the coordinates of the pupil and iris coarse contours) are given by the segmentation step to use during the normalization step. Each point of the contour has coordinates (x, y, φ), where (x, y) is the coordinate of the radius relative to the estimated centre and φ the angle between 0 to 2π. These contours

parameters are a required input for the normalization step. Figure 3(a) shows an example segmented iris image and Figure 3(b) shows an example binary mask image.



(a) Segmented image (b) Binary mask image

Figure 3. An example of iris segmentation result

2.3.2 Normalization

In the normalization stage, the iris texture is transformed into a fixed-size rectangular block. After finding the centre of pupil and the inner and outer boundaries of iris, the texture is transformed from Cartesian to polar coordinates with fixed radial and angular resolutions. The normalization process produces iris regions which have the same constant dimensions so that two images of the same iris under different conditions will have same features at the same spatial location.

The Daugman's rubber sheet model is applied for normalization. Let W and H be the width and height of the desired normalized image respectively. A regular sampling of angles θ_k is computed as in equation (1).

$$\theta_k = 2\pi \frac{k}{W} \tag{1}$$

where k ranges from 0 to W, $\theta_0 = 0$ and $\theta_w = 2\pi$.

Let x_p, y_p, φ_p and x_i, y_i, φ_i be the coordinates of the pupil and iris boundaries where $(x\ , y)$ are the x -coordinate and y -coordinate of the radius relative to the estimated centre and φ is the angle of non-regular sampling. The point (X_k^p, Y_k^p) which corresponds to θ_k is estimated by interpolating the two nearest points j and $j+1$ to θ_k as in equation (2). Figure 4 shows the computation of (X_k^p, Y_k^p) coordinate point.

$$X_k^p = (1-\alpha) * x_p^j + \alpha * x_p^{j+1} \tag{2}$$

$$Y_k^p = (1-\alpha) * y_p^j + \alpha * y_p^{j+1} \tag{3}$$

$$\alpha = \frac{\theta_k - \varphi_p^j}{\varphi_p^{j+1} - \varphi_p^j} \tag{4}$$

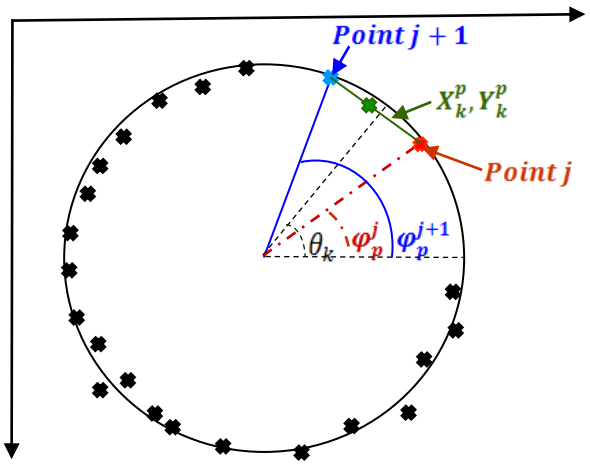
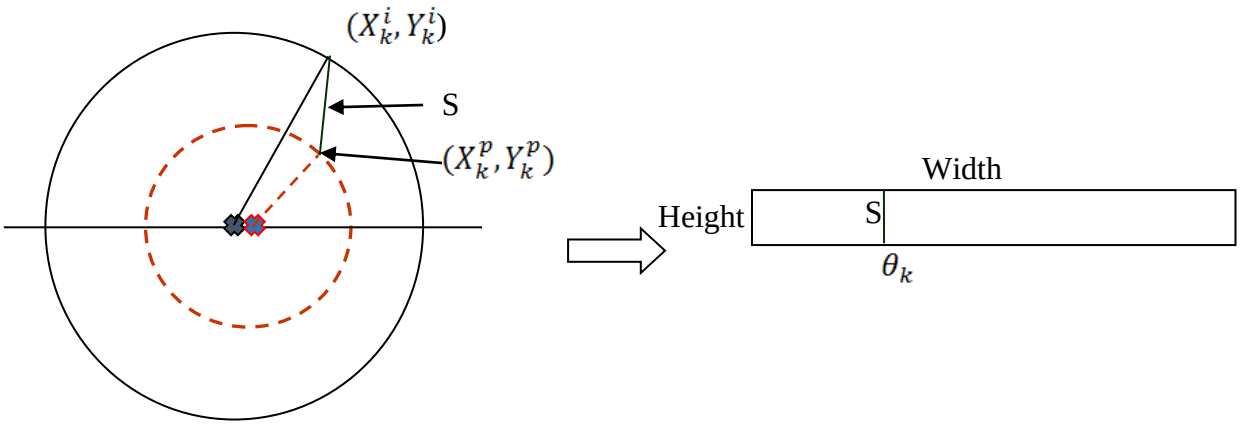


Figure 4. The Computation of the coordinate point (X_k^p, Y_k^p)

The point (X_k^i, Y_k^i) of the iris contour is computed in a similar way. Let S be the segment between $(X_k^p, Y_k^p)$and (X_k^i, Y_k^i). Then S is rescaled to fit with the height H of the normalized image as shown in Figure 5(a) and (b).



(a) The segment S between pupil and iris coordinates (b) The normalized image

Figure 5. The rubber sheet model to transform circular iris into fixed dimension

Although the rubber sheet model accounts for pupil dilation, imaging distance and non-concentric pupil displacement, it does not compensate for rotational inconsistencies. In the Daugman's rubber sheet model, rotation is accounted for during matching by shifting the iris templates in the θ direction until two iris templates are aligned.

The normalization step also transforms the mask into a normalized mask. The normalized mask is not required for the encoding step but is an optional input of the matching stage. Figure 6 (a) and (b) shows an example of normalized image and normalized mask respectively.

(a) A normalized iris Image

(b) A normalized mask

Figure 6. An example of iris image normalization

2.3.3 Feature Encoding

Feature encoding aims at creating a template containing only the most discriminating features of the iris. In the encoding stage, features such as iris textures are extracted by a bank of the Gabor filters which is customizable in orientation and resolution. The normalized image is filtered by six Gabor filters, each of which has two independent real filters, real part and imaginary part, and the result is thresholded relative to zero in order to form binary image. So the result of each pixel for Gabor filter is encoded on two bits (real part and imaginary part as positive and negative). The whole binary image of the size W × (n×H) where n is the number of Gabor Filters and W × H is the size of normalized image, is outputted as the iris code or iris template. Figure 7 shows an example of iris template image.

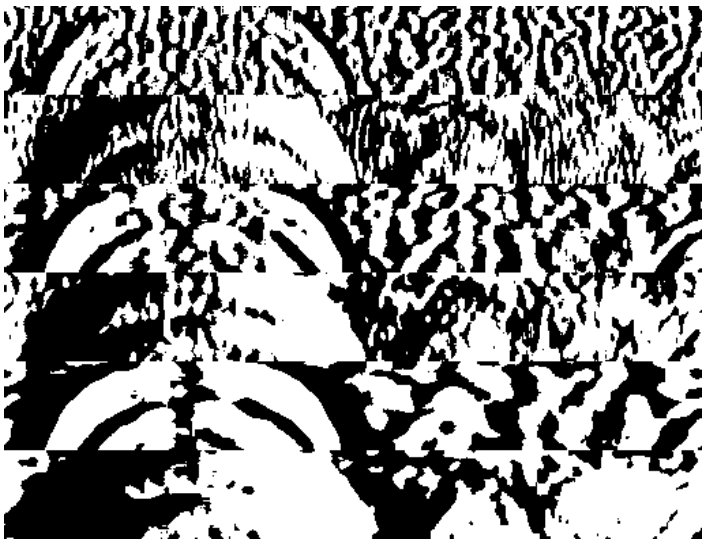


Figure 7. An example of iris template image

2.3.4 Matching

In the matching stage, two eyes are compared using the Hamming distance between their iris codes. Not all pixels are used for matching and they are indicated by the matrix of application points chosen within the iris template. Firstly, a total mask is built to ignore noisy pixel, that is the result of logical *and* between the matrix of application points and the masks. Then logical *xor* between both iris codes is computed for the application points that are not masked.

The Hamming distance gives a measure of how many bits are the same between two bit patterns. A decision can be made based on the hamming distance between two bit patterns. When we compare the iris codes X and Y, the Hamming distance(HD) is calculated in the following equation (5).

$$HD = \frac{1}{N}\sum_{j=1}^{N} X_j (XOR) Y_j \tag{5}$$

where HD is the sum of disagreeing bits (sum of the exclusive-OR between X and Y) over N, where N is the total number of bits in the bit pattern. Table 2 shows the truth table for XOR operation. If the bits in two patterns are same in XOR operation, it will produce 0 as the result bit and if the bits in two patterns are not same, it will produce 1 as the result bit. The result bit pattern is obtained by performing XOR operation on the bit pattern of X and Y iris codes. If two patterns are derived from the same iris, the Hamming distance between them will be close to 0.0, since they are highly correlated and the bits agree between the two iris codes.

Table 2. Truth table for Exclusive OR operation

X bit	Y bit	XOR
0	0	0
0	1	1
1	0	1
1	1	0

CHAPTER 3. EYE REGION DETECTION BY LIKELIHOOD COMBINATION

3.1 Problems in Previous Works

The iris image acquisition condition plays as one of the important roles in iris authentication system. Most commercial iris acquisition systems work in near infrared (NIR) band. The conventional system has been very restrictive in acquisition process. Users have to fully cooperate like standing close to the camera with a specified distance and looking straight at it. The eye has to be wide open and located in the centre of the image. The specular reflections should appear inside the region of the pupil and not on the pupillary boundary or inside the iris region. In less constrained environment where many constraints are impossible to undertake on the users, the degree of freedom of subjects' movement and position increase. So, the images taken under such environment increase variations a lot. These variations accelerate differences in illumination, eye pose, resolution in different acquisitions. The similarity between two irises of the same eye decreases when the acquisition environment differs, which lead to confuse the biometric decision whether the same person or not.

OSIRIS can work well for images that have taken focus on the iris region under clear background without glasses and illumination [15]. The segmentation methods generally find one best iris candidate and segment the iris region. As the segmentation stage of OSIRIS depends on the pupil centre and coarse contours, it is possible to cause error if it detects the location of pupil centre wrongly. When OSIRIS is applied on the CASIA-v4-Distance, the segmentation module performs wrongly on the other organs of the face such as nostrils, mouth, glasses and hair. So, we add an eye region detection stage before segmentation. The eye region is firstly detected by our proposed likelihood combination method and giving a detected eye image as input image for OSIRIS can get better results than simple face input images as it doesn't need to search the iris on all other organs of the face.

3.2 Our Proposed Eye Region Detection by Likelihood Combination

Eye detection is a pre-requisite stage for many applications such as human–computer interfaces, iris recognition, driver drowsiness detection, security and biology systems. In this work, we introduce a new eye region detection method for iris authentication.

In our proposed method, likelihood combination is applied to detect an eye region from a given input face image. For each image, feature points from the half face image are extracted by AKAZE [3] feature matching and likelihood is created using these feature points. At the same time, the matched points are extracted by template matching from the same half face image and likelihood is created using these matched points. Both likelihoods from AKAZE and template matching are combined and the location of eye is determined.

AKAZE feature matching and template matching tend to find different position on the face image. AKAZE feature matching tends to find feature points on nostrils, nose pad and pad arm of eyeglasses, near lip line and iris, while template matching tends to find matching points on the forehead, nose, near rim of eyeglasses, eyebrows and iris. So, these two methods are incorporated to get better performance for eye region detection.

The following Figure 8 shows the overview of our proposed likelihood combination method for eye region detection. We introduce likelihood images to determine the location of eye region. The pixel value of the likelihood image indicates the likelihood at that position. The highest pixel value in the likelihood image means the eye region exists in that location. For most images, these two methods can locate the actual eye position as the first or the second peak in the likelihood images. Even when both likelihood images from these two methods indicate the actual eye as the second peak, combining these two second highest peaks can be the highest peak in the combined likelihood image. Thus, the combined likelihood peak can indicate the accurate eye location. Based on the highest peak location of the combined likelihood image, the input face image is cut only around the eye area to be processed in OSIRIS.

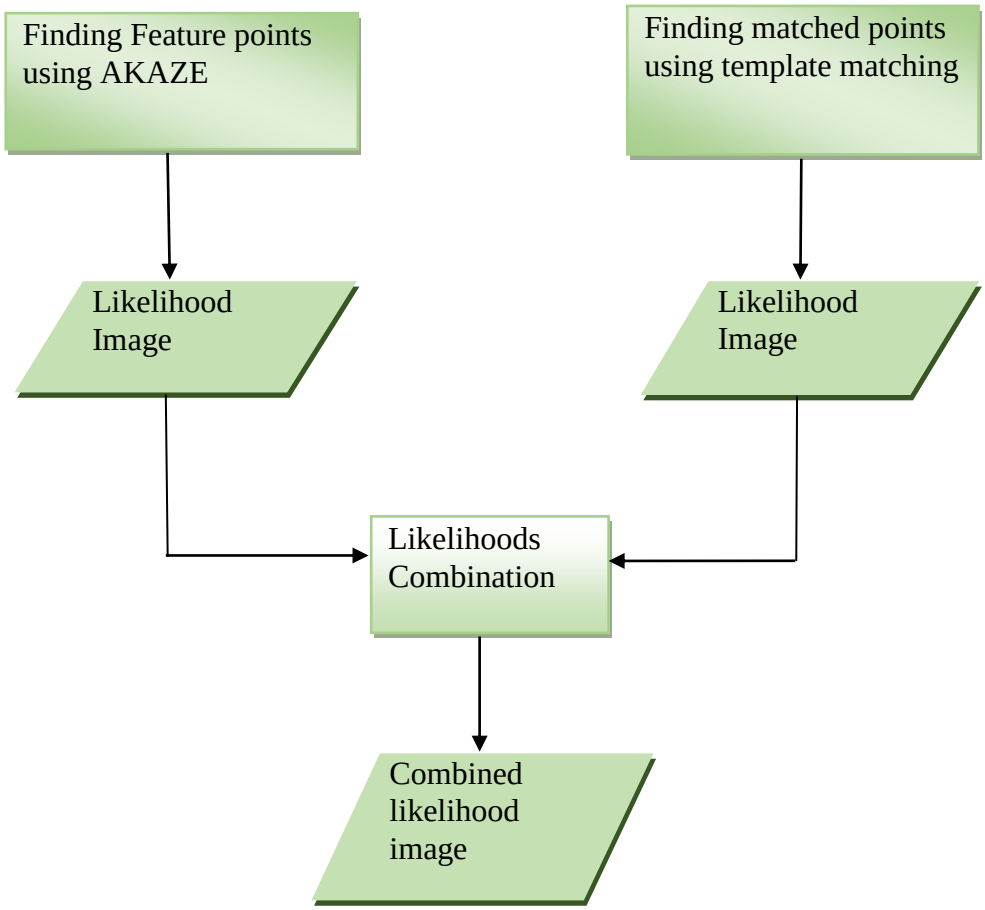


Figure 8. The diagram of likelihood image combination

3.3 Related Works

In the literature review, there are a lot of eye region detection methods. Zhu et al. [16] broadly classified existing eye detection methods into two categories: the active infrared (IR)-based approaches and the traditional image-based passive approaches. In the active infrared-based approaches, the eye regions are effectively detected and tracked by utilizing the principle of red-eye effect in flash photograph with special IR illuminator for image acquisition. Although it is relatively simple and effective, it requires a stable lighting condition and a camera close to the subject.

The image-based passive approaches such as template matching method [17, 18], eigenspace method [19] and hough transfom method [20] utilize specific properties of eye such as intensity and shape.

The template matching method compares the given template image with an input image and finds the most similar sub-region in the input image. It cannot work well on images in rotation and illumination. So, multiple templates are used to find the eye region. Yuille et al. [21] proposed a deformable template for face feature extraction. An eye was described by a parameterized template and an energy function was defined to link edges, peaks, and valleys in the input image to corresponding parameters in the template. The advantage of this approach is that besides the location, more features of an eye, such as its shape and size, can be obtained at the same time. But the approach is time consuming, and its rate of success relies greatly on the initial position of the template.

Pentland et al. [19] used an eigenspace method for eye detection. This method can achieve better eye detection performance than a simple template matching method since training samples have covered different eye variations in appearance, orientation and lighting conditions. One problem with this method is that its performance is closely related to the selection of the training set. Another drawback is that it requires the training and test images to be normalized in size and the training and testing images of various orientation.

The hough transform method is based on the shape of iris and is often applied to binary edge maps. The drawback of this approach is that the performance depends on threshold values used for binary conversion.

Besides these three classical approaches, recently many other image-based eye detection techniques have been reported. In Guo Can Feng [22] work, they present an eye detection method for grey intensity image. The proposed method uses multi-cues for detecting eye windows from a face image. Three cues from the face image are used, and a

cross-validation process is performed based on these cues. This process generates a list of possible eye window pairs. For each possible case, variance projection function is used for eye detection and verification. In Zhou and Geng [23] work, the eye region is detected based on the intensity of the eye area. They applied the hybrid projection function (HPF) that uses two characteristics of eye area obtained by individually exploiting the variance projection function and the generalized projection function. Kawaguchi and Rizon [24] located the iris using intensity and edge information. The main techniques included in their algorithm are a feature template, a separability filter, Hough transform and template matching. Sirohey and Rosenfeld [25] proposed an eye detection algorithm based on linear and nonlinear filters. In Huang and Wechsler's method [26], the task of eye location was considered as a test bed for developing navigation routines implemented as visual routines. They used genetic algorithms and built decision trees to detect eyes. For the purpose of face detection, Wu and Zhou [27] employed size and intensity information to find eye-analog segments from a grey scale image, and exploited the special geometrical relationship to filter out the possible eye-analog pairs. Similarly, Han et al. [28] applied such techniques as morphological closing, conditional dilation and labelling process to detect eye-analog segments. Hsu et al. [29] used colour information for eye detection.

CHAPTER 4. LIKELIHOOD COBINATION METHOD

4.1 Likelihood Creation by AKAZE Feature Matching Method

In this research, we apply AKAZE [3] feature matching to create likelihood image. AKAZE (Accelerated-KAZE) is an improved version of KAZE [30] feature detection algorithm. AKAZE feature matching is now popular in a wide range of applications in computer vision such as real time feature extraction for object recognition and 3D reconstruction. AKAZE feature detection contains three major stages: building non-linear scale space, feature detection and binary descriptor generation. Firstly, fast explicit diffusion(FED) scheme is used to build non-linear scale space to detect and describe features. The key points are located in the second stage. The determinant of the normalized Hessian matrix for each of the filtered images is computed in the nonlinear scale space, and the local maxima at each sublevel are picked out as candidate key points. In the third stage, AKAZE introduces a modified-local difference binary (M-LDB) descriptor that exploits gradient and intensity information from the nonlinear scale space. M-LDB estimates the main orientation of the key points based on the histogram. The grid of LDB is rotated according to the direction of the key point. Finally, the binary descriptors are generated through comparison among the grids. So, AKAZE features are highly robust to rotation and shows good performance on low-texture objects and people.

In our research, an eye image is given as a sample image and a half face image is given as an input image to detect the eye region. AKAZE features are extracted from the input and sample image, and descriptors for each features are computed. The descriptors from each image are checked bi-directionally to extract correlated feature points. Brute force matching which compares a single feature in one image with all other feature descriptors in other image is used to find the corresponding features between each image and returns a pair of features with the minimum difference. The brute force matching is exploited in a greedy manner to obtain multiple correlated features. These final correlated features are used to create a likelihood image.

An example of the correlated feature points by AKAZE feature matching are shown in Figure 9. The final correlated features are shown in Figure 10 (a). Plural feature points are extracted as the correlated feature points and the likelihood image created by using AKAZE feature points are shown in Figure 10 (b).

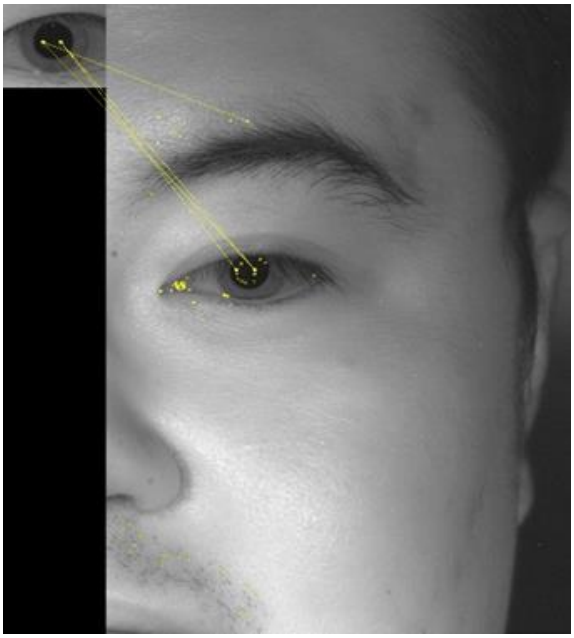


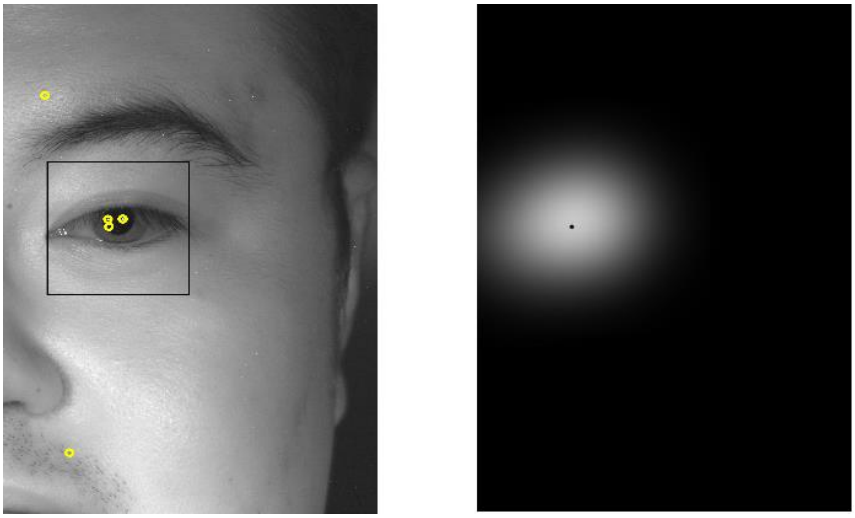
Figure 9. Feature points detection by AKAZE

The feature points obtained from AKAZE feature matching are used to create the likelihood image. Based on the correlated feature points, the likelihood of eye is calculated at each position of the likelihood image. It is high possibility where many correlated feature points exist is the location of the eye. In order to increase the likelihood of the eye location where many feature points exist, the likelihoods are calculated by voting around the feature points. To create a likelihood image, an image having the same size as the input image is prepared firstly, and the pixel values are initialized to 0. The pixel value in the likelihood image is voted using the Gaussian distribution weight around correlated feature points. Let F = $\{\boldsymbol{f}_i\}$ be the set of the coordinate of correlated feature points. For each feature point, we vote the value calculated by equation (6) to the neighbour pixels at $\boldsymbol{x}$ in the likelihood image.

$$L(\boldsymbol{x}) = exp\left[\frac{(\boldsymbol{x}-\boldsymbol{f}_i)^2}{2\sigma^2}\right] \tag{6}$$

where σ is the standard deviation, $\boldsymbol{x}$ is the pixel coordinate respectively. After voting for all correlated feature points, the image is normalized so that the maximum pixel value becomes 1. As the result, the likelihood increases at positions where many feature points

gather. An example likelihood image is shown in Figure 10 (b), the white pixels show the higher likelihood of the eye location. The black spot within the white pixel region indicates the highest peak in the likelihood image.



(a) AKAZE correlated feature points (b) The likelihood image by AKAZE

Figure 10. AKAZE feature matching and likelihood creation

4.2 Likelihood Creation by Template Matching Method

In this research, we also use template matching to create likelihood image. Template matching is widely used in many computer vision and image processing applications such as facial recognition and eye detection. A sample sub-image is used to find the most similar object in the input image.

Template matching finds the sub-image from the input image which is most similar to the template image. In this research, template matching is used to find the location of eye region from a face image. The position of eye is determined by a pixel-wise comparison of the face image with a given eye template image. The template is moved to x-direction for *u* times and y-direction for *v* times in a large input image and compared how similar the template and input images are for each (*u*, *v*) position.

The similarity value at each position is computed by Normalized Correlation Coefficient (Normalized Cross-Correlation) which is robust for lighting condition changes, as in equation (7).

$$R(x,y) = \frac{\sum_{x,y}(t(x,y)-t')(f(x+u,y+v)-f')}{\sqrt{\sum_{x,y}(t(x,y)-t')^2 \sum_{x,y}(f(x+u,y+v)-f')^2}} \quad (7)$$

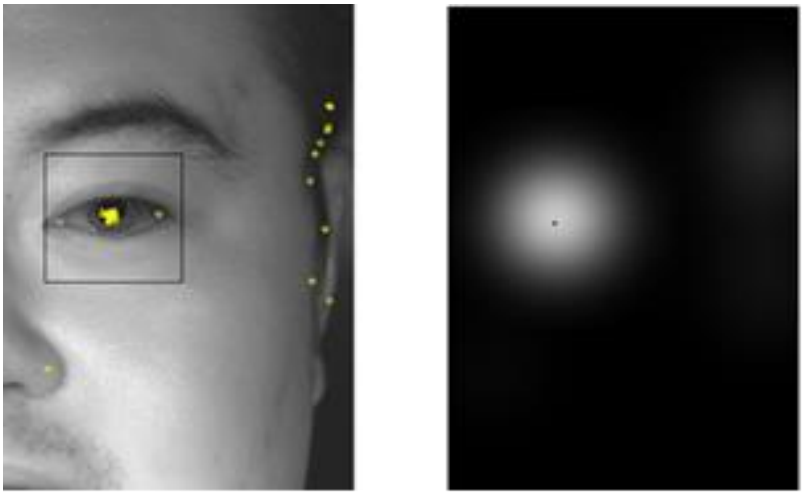
$$t' = \sum_{(x,y)} t(x,y)/k^2$$

$$f' = \sum_{(x,y)} f(u+x, v+y)/k^2$$

where $f(x,y)$ and $t(x,y)$ are the pixel values at (x,y) in the input image and in the template image respectively, and k^2 is the size of template image. $R(x,y)$ returns the maximum similarity value as '1'. Correlation is the measure of agreement between two corresponding pixel values of template and input images. The more correlated, the higher value will be.

The matched location of the eye template image is the position which has largest similarity value in the input image. A number of template images(N_T) are used to find the corresponding eye region for each face image, and one matched point from each template image is extracted by template matching. The number of matched points is same as the number of template images.

The likelihood image is created by all matched points in the same manner as AKAZE feature points. There is a high possibility where many matched points exist in the likelihood image is the location of our eye region. Figure 11(a) shows the template matching matched points and Figure 11(b) shows an example of likelihood image created by template matching matched points.



(a) Matched points (b) The likelihood image

Figure 11. Template matching method and likelihood creation

4.3 Likelihood Combination

The likelihood images generated from AKAZE feature matching and template matching are simply added and a combined likelihood image is obtained. The combined likelihood image is normalized so that the maximum pixel value will be 1. The advantage of adding likelihood images is that the sum of two second peaks can generate the highest peak in the combined likelihood image.

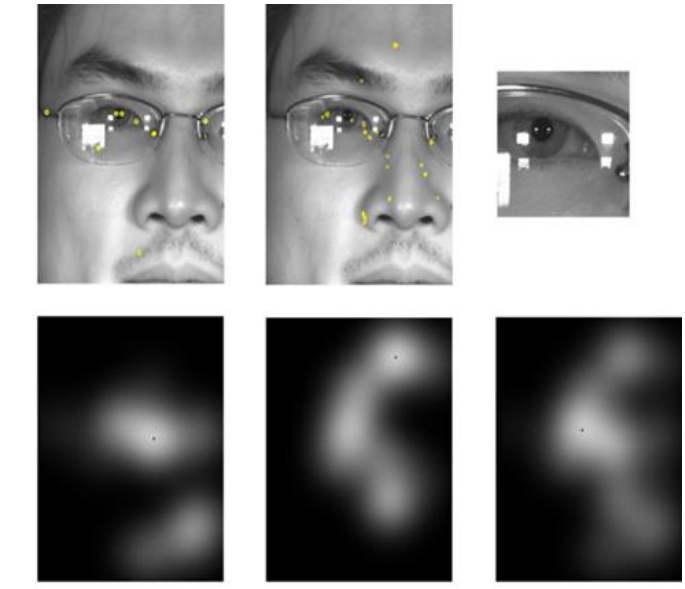
Let $I_a(x, y)$ be the likelihood image obtained from AKAZE feature matching and $I_t(x, y)$ be the likelihood image obtained from template matching at (x, y) coordinates. The image $I_r(x, y)$ be the result combined likelihood image. The pixel values at position (x, y) from two likelihood images (I_a and I_t) are added and saved as the combined likelihood image I_r as in equation (8).

$$I_r(x, y) = I_a(x, y) + I_t(x, y) \quad (8)$$

The proposed likelihood combination method outperforms AKAZE and template matching methods. It is because each method finds the different locations of matched points. As an example, a half face image is given as an input image and the eye region is detected by AKAZE and template matching. The detected feature points by each method are shown in Table 3. We can see that the feature points detected by AKAZE and template matching are located at different positions. Therefore, the combined likelihood can generate the correct eye location since matched feature points locations from AKAZE and template matching are same in iris. So, likelihood combination can output the accurate eye region for regions that are not detected accurately by each method. An example likelihood combination and cut image is shown in Figure 12.

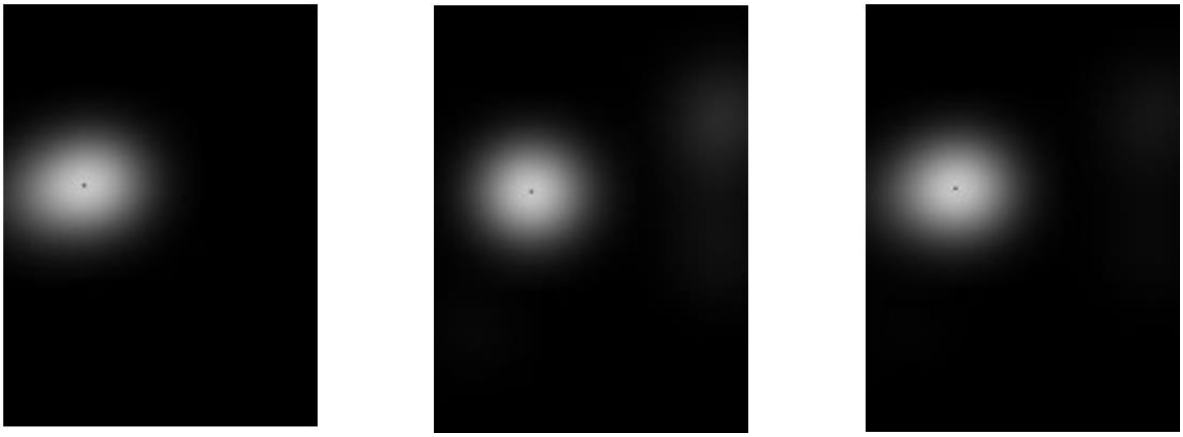
Table 3. Location of key points at different regions detected by AKAZE and template matching

Regions	AKAZE feature points	Template matched points
Near rim	-	3
Nose pad	4	-
Pad arm	5	-
Nostril	4	-
Nose	-	12
Vertical lip lines	6	-
Eyebrows	-	4
Forehead	-	13
Eyeglasses	-	3
Iris	13	5
Total	31	40



(a) AKAZE (b) Template Matching (c) Combined

Figure 12. Feature points and likelihood images

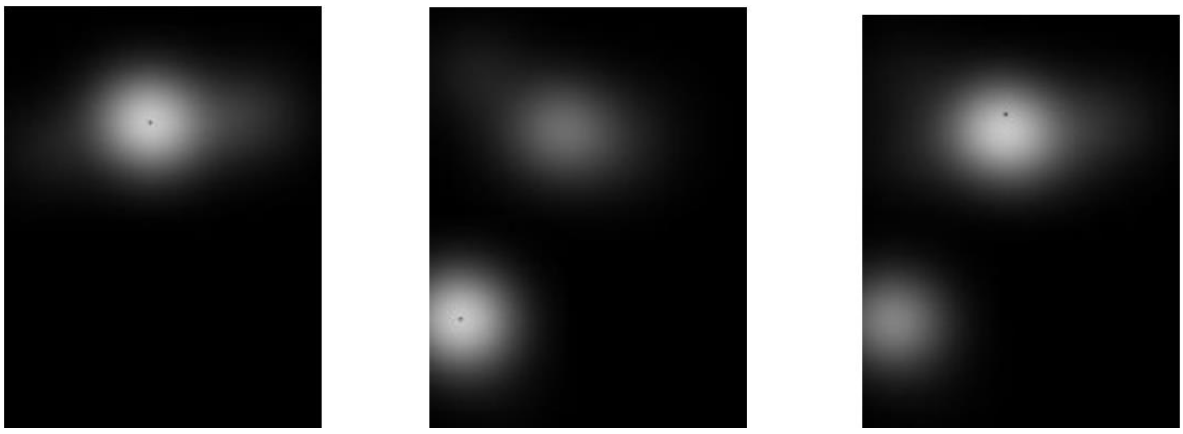


(a) AKAZE (b) Template Matching (c) Combined

Figure 13. The likelihood images
(AKAZE: correct, Template matching: correct, Combined: correct)

Figure 13, 14, 15 and 16 show the likelihood images created (a)using AKAZE feature points, (b) using template matching matched points and (c) by combining two likelihood images from AKAZE and template matching. The black spot within the white region in the images indicates the correct eye position.

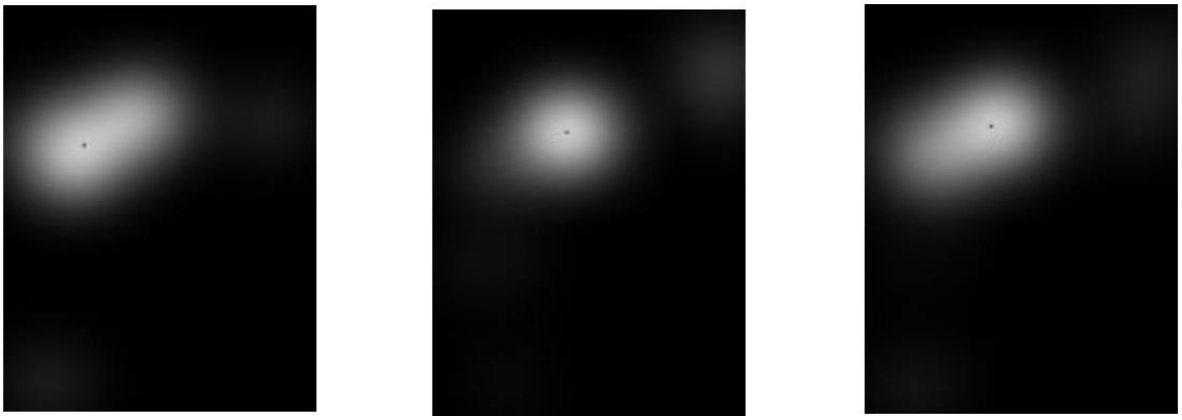
In Figure 13, since both likelihood images obtained from AKAZE and template matching correctly indicate the eye region as the highest peak, then combining two highest peaks results the highest peak in the combined likelihood image. So, the correct eye region is obtained from the combined likelihood image.



(a) AKAZE (b) Template Matching (c) Combined

Figure 14. The likelihood images
(AKAZE: correct, Template Matching: wrong, Combined: correct)

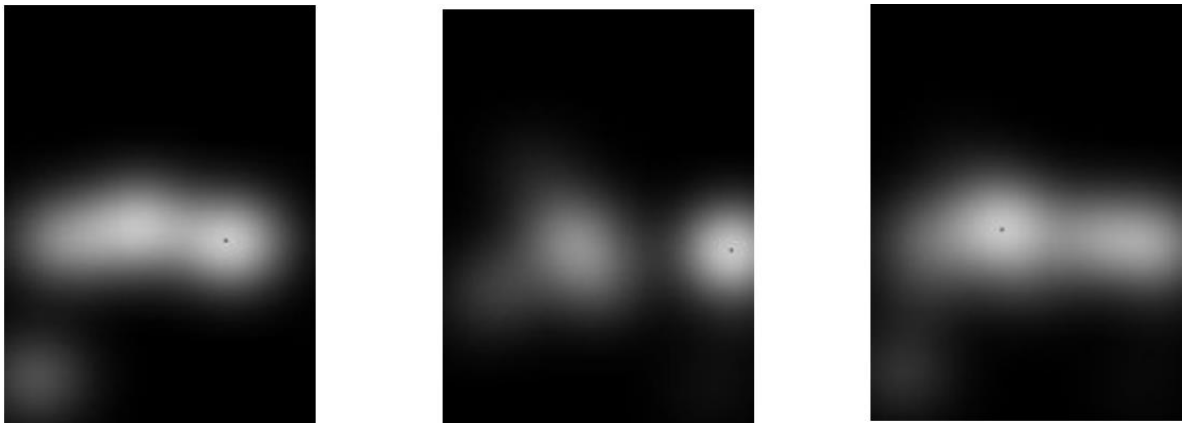
In Figure 14, the likelihood image obtained from AKAZE correctly indicates the eye region as the highest peak, while the likelihood image obtained from template matching doesn't indicate the eye region as the highest peak. Contrary, in Figure 15, the likelihood image obtained from AKAZE doesn't indicate the eye region as the highest peak, while the likelihood image obtained from template matching correctly indicate the eye region as the highest peak. In both cases, combining highest peak and second highest peak results the highest peak in the combined likelihood image. So, the correct eye region is obtained from the combined likelihood image.



(a) AKAZE (b) Template Matching (c) Combined

Figure 15. The likelihood images

(AKAZE: wrong, Template Matching: correct, Combined: correct)



(a) AKAZE (b) Template Matching (c) Combined

Figure 16. The likelihood images

(AKAZE: wrong, Template Matching: wrong, Combined: correct)

In Figure 16, both likelihood images obtained from AKAZE and template matching don’t indicate the eye region as the highest peak, but as the second highest peak. Combining two second highest peaks results the highest peak in the combined likelihood image. So, the correct eye region is obtained from the combined likelihood image even in this case.

CHAPTER 5. EXPERIMENTAL ANALYSIS AND PERFORMANCE ASSESSMENT

5.1 CASIA-v4-Distance Dataset

CASIA-v4-Distance [4] is a dataset captured at a distance under NIR (near infrared) wavelength. The images from this database were acquired under less-constrained environment. CASIA-v4-Distance contains iris images captured using the self-developed long-range multi-modal biometric image acquisition and recognition system. The stand off distance is around 3 meters. The images include most of facial features and iris patterns. There are 2567 face images of 142 persons that are not rotated and iris size is almost the same for all images. There are 9~20 face images for each person. The images have relatively lower quality. Although they are capture under NIR wavelength, since at a distance, the iris region has low resolution. The iris texture is visually less clear and the noise in this dataset is relatively heavier. Eye glasses, specular reflection on the eye glasses, and blurred images are included in these face images.

In this research, we firstly divided the original images into the left and right half face images as the left and right irises are not the same for each person. Then, the left and right eye regions are detected using our proposed likelihood combination method. The following Figure 16 shows an example image of CASIA-v4-Distance image dataset. Each image has the same size of (2352,1728) width and height.

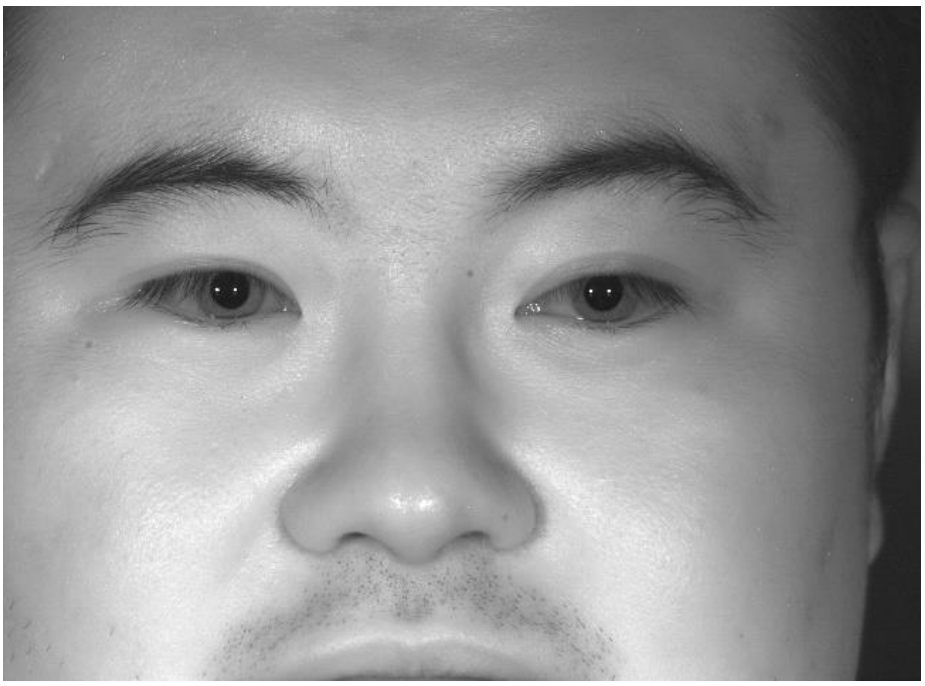


Figure 17. A sample image in CASIA-v4-Distance

5.2 Eye Region Detection Results

The performance of the proposed approach has been evaluated on the public data set of CASIA-v4-Distance [4]. For 2567 left and 2567 right images, the eye regions are detected and cut by our proposed method and used as input images for OSIRIS.

For AKAZE feature matching, one eye image from 2567 left and 2567 right images respectively is selected as a sample image. The following Figure 18(a) and (b) shows the left and right sample images respectively used in the AKAZE feature matching. The original left image S4001D00L.jpg is cropped at position (476,788) and (680,918) and the original right image S4001D06R.jpg is cropped at position (666, 405) and (837,543).

(a) S4001D00L.jpg (b) S4001D06R.jpg

Figure 18. The left and right sample eye images used in AKAZE

For template matching, $N_T = 40$ eye images for left and $N_T = 30$ images for right are chosen as template images. Some blurred images are chosen as template images. Figure 19 shows the template images for left eye region detection and Figure 20 shows the template images for right eye region detection. For likelihood creation, the value of σ is set σ = 200 in Gaussian distribution function to vote each position in the likelihood image.

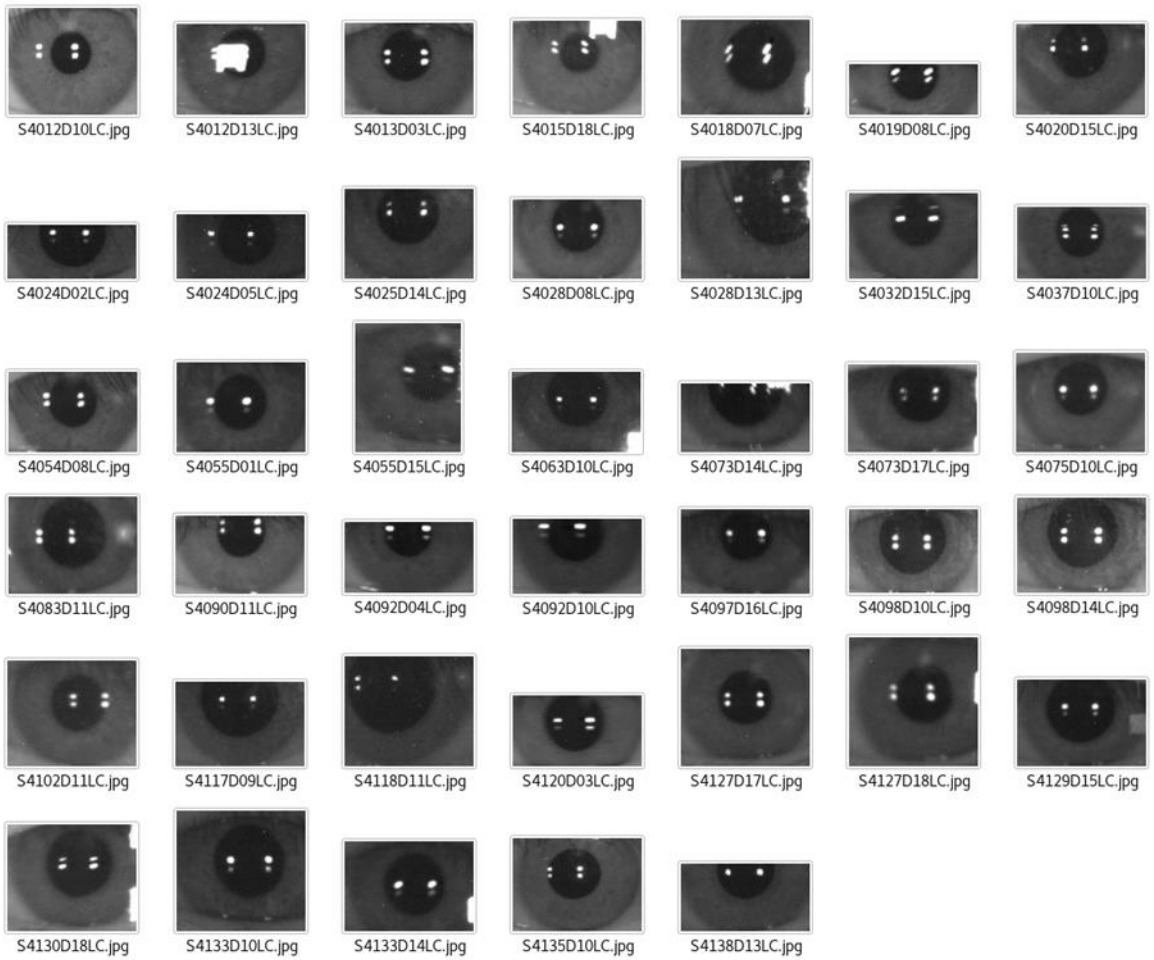


Figure 19. The template images used for left eye detection

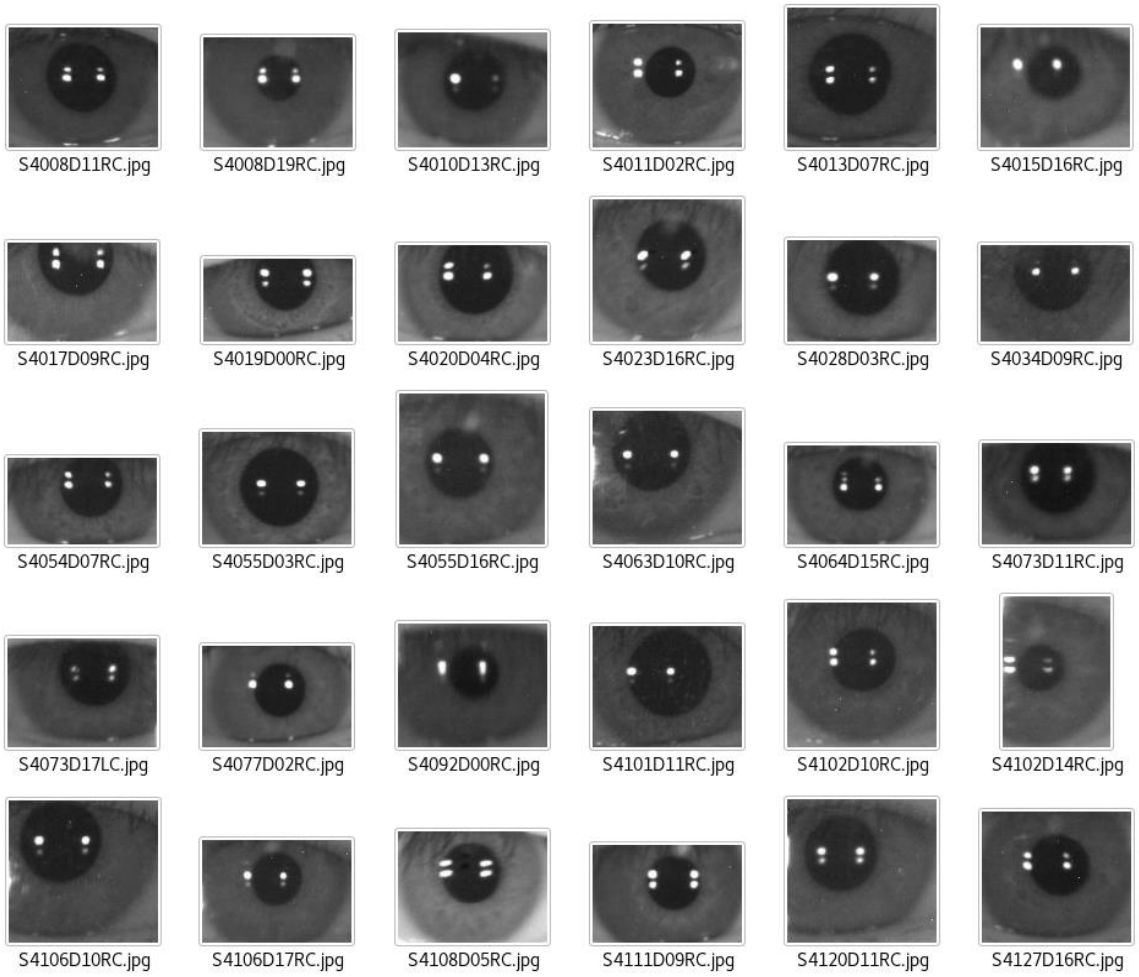


Figure 20. The template images used for right eye detection

The eye region is cut based on the maximum location in the combined likelihood image. The maximum location is set as the centre point and a rectangle area of the fixed size (400 × 400) is cut as the eye region. We assume iris size is almost same in the given input images. A sample cut image is shown in Figure 21. The results for detected eye region images are counted manually as correct if the iris region is fully included in the cut image and as wrong if any part of iris region is not included in the cut image.

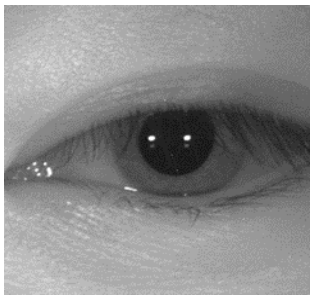


Figure 21. The cut image from likelihood combination

Table 4 shows the cut image results by AKAZE, template matching and likelihood combination with the wrongly detected region in the face. The specular reflection on the eye region especially who wear glasses causes the images to be wrongly detected and cut on other organs and parts such as pad arm, end piece, bridge, mouth, eyebrows, hair, nose and sclera. The image with detected eye region is given as the input image to the OSIRIS.

Table 4. Checked cut images' results
by AKAZE, template matching and likelihood combination

Wrong places	Pad arm	End piece	bridge	mouth	Eye brows	Hair	Nose	Sclera	Total
Wrong by AKAZE	127	16	3	1	4		1	1	143
Correct by AKAZE									4991
Wrong by Template matching	95	24		21	4	17	20	12	193
Correct by Template matching									4939
Wrong by Combine	68			1			3	8	80
Correct by Combine									5054

Table 5. Eye region detection results
by OSIRIS, AKAZE, template matching and likelihood combination

	Correctly detected
OSIRIS	4867/5134(94.7%)
AKAZE	4991/5134 (97.2%)
Template	4939/5134 (96.2%)
Combined	5054/5134 (98.4%)

Table 5 and Figure 22 show eye region detection results by OSIRIS, AKAZE, template matching and likelihood combination methods. Although original OSIRIS doesn't detect eye region, to be more prominent we calculate the eye region detection for OSIRIS. As OSIRIS firstly locate pupil area to segment the iris region, we check manually whether the detected pupil circle is within the eye region which is the same size of the eye region cut by other methods. If the pupil circle is within the size of eye region, it is noted as correct one and otherwise as wrong one.

AKAZE can detect eye region for blurred images and for some images with pupil dilated or constricted due to lighting condition. Template matching can detect eye region for images with normal pupil and iris sizes and clear images. But both methods cannot work well for images with severe illumination caused by specular reflection especially on the eye surface. OSIRIS can work for clear images and it cannot work well for blurred images and images with noise such as illumination. Likelihood combination can detect eye region more correctly than other two methods.

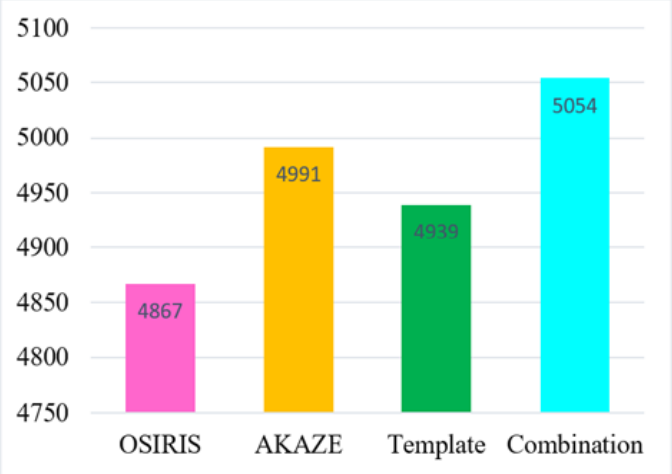


Figure 22. Correct eye region detection results

5.3 Iris Segmentation Results

Next, we continue iris segmentation for half face images and eye region images detected by likelihood combination. Table 6 shows the manually checked iris segmentation results for detected eye images by likelihood combination, and the results of the half face images segmented by OSIRIS.

The segmented iris images are regarded as correct if the iris and pupil are accurately segmented as two circles: the inner circle between pupil and iris contour and the outer circle between iris and sclera contour. If the segmented iris circle contains inaccurate iris contour or a little bigger circle contour than normal iris contours, it is regarded as wrong.

Giving images cut by likelihood combination as input to OSIRIS can get better segmentation results than simple face input images as it can avoid to segment wrongly other organs on the face.

Table 6. Iris segmentation results

	Correctly Segmented
OSIRIS	4389/5134(85.5%)
Combined	4502/5134(87.7%)

5.4 Overall System Outcome and Analysis

We conduct an iris authentication experiment to calculate the overall accuracy. Firstly, 2567 face images from CASIA-v4-Distance are cropped into two half faces in order to detect left and right eyes respectively. As the left and right iris of the one person is not same, we get totally 5134 (2567 × 2) half face images from 2567 full face images. Here, we check the authentication results by giving two kinds of input images to OSIRIS, one kind of image is simple half face image and another kind is cut eye region image that is obtained by using our likelihood combination method. All input images are segmented, normalized, encoded at first and then matched in matching stage in OSIRIS.

As an experiment, we use leave-one-out test in order to check the authentication results. It means that we choose one image from all images in the dataset as one query image and all other images except the query image is let as gallery images. As an example, if we choose S4000D00L.jpg as a query image, we will leave this image when counting the gallery images. The Hamming distance between iris codes of this query image (S4000D00L.jpg) and all gallery images are computed in the matching stage and the minimum distance is searched between this query and other gallery images. If the minimum Hamming distance between query and gallery images of the same person is found, we regard this query image as the correctly authenticated image. After one query image have finished, it is the time for next image to be placed as query image. We continually perform the same procedure until all left and right images are chosen as query images for authentication. In this point, there are some images that OSIRIS cannot work because of so bad input images. In 5134 half face images, OSIRIS cannot work for 21 images. In 5134 cut eye region images, OSIRIS cannot work for 17 images. So, we use only 5113 half face images and 5117 cut eye region images for matching stage. We repeatedly do the leave-one-out test until 5113 half face images are chosen as query images and 5117 cut eye region images are chosen as query images. Table 7 shows the authentication result of this leave-one-out experiment. Here, there is a case that different images of a same person are segmented wrongly and the wrongly segmented regions are very similar. For example, the same portion of the nose in the different images for the same person is wrongly segmented especially when the person wears glasses. At that time, wrongly segmented query image and wrongly segmented gallery image of the same person match each other with the minimum distance as the segmented portions are same for this person. As a drawback, the authentication results are nearly same for OSIRIS and giving detected eye image to OSIRIS.

Table 7. Iris authentication results of leave-one-out test

	Correctly Authenticated
OSIRIS	4684/5113(91.6%)
Combined	4708/5117(92%)

Next, we make another experiment to get correct and prominent authentication results. We choose correctly segmented four images in both methods for each person from 5113 half face images and 5117 cut eye region images to make gallery images. Since each person has left and right eye images, we get total 1136 gallery images (4×142 persons=568 left gallery images and 4×142 persons=568 right gallery images). Other images that are not used as gallery images are set as query images. Now, we have 3977(5113-1136) query images for half face images and 3981(5117-1136) query images for cut eye region images. The minimum Hamming distance between iris codes of query and gallery images is calculated in the OSIRIS matching stage. If the minimum distance between query and gallery image is found from the same person's eye image, it is regarded as correctly authenticated one, else it is wrong. The overall system accuracy is calculated as the fraction of the number of correctly authenticated query images divided by total number of query images.

Before checking all query images detected by OSIRIS and our likelihood combination method, we check the correctly detected images by both methods. Table 8 shows the iris authentication results for correctly detected images by both OSIRIS and likelihood combination methods. Both methods show almost same performance in this case. For correctly detected images, the result is nearly same as both methods can detect eye region accurately.

Table 8. Iris authentication results for images correctly detected by both methods

	Correctly Authenticated images/ (correctly segmented images - gallery images)
OSIRIS	3062/3253(94%)
Combined	3052/3366(90.7%)

Next, we check for wrongly detected images by OSIRIS to clearly see that giving cut eye region images to OSIRIS shows better performance than giving simple half face input image for the authentication result. Figure 23 shows the authentication results for images wrongly detected by OSIRIS. Out of 247 images wrongly detected by OSIRIS images, likelihood combination method can detect correct eye region for 195 images

(78.9%) and segment iris area for 117 images (47.4%) and authenticate 111 images (44.9%) correctly. So, our proposed method can detect eye region very well for face images captured under less constrained environment. In such a way, detection of eye region before iris segmentation can avoid segmentation on other organs of the face. Furthermore, once we can segment iris correctly, the authentication results are very well (111/117 = 94.9%). However, iris segmentation result by OSIRIS is not so good. If we can segment all detected images accurately, the performance of the system will be higher. The improvement of iris segmentation method is left as a future work.

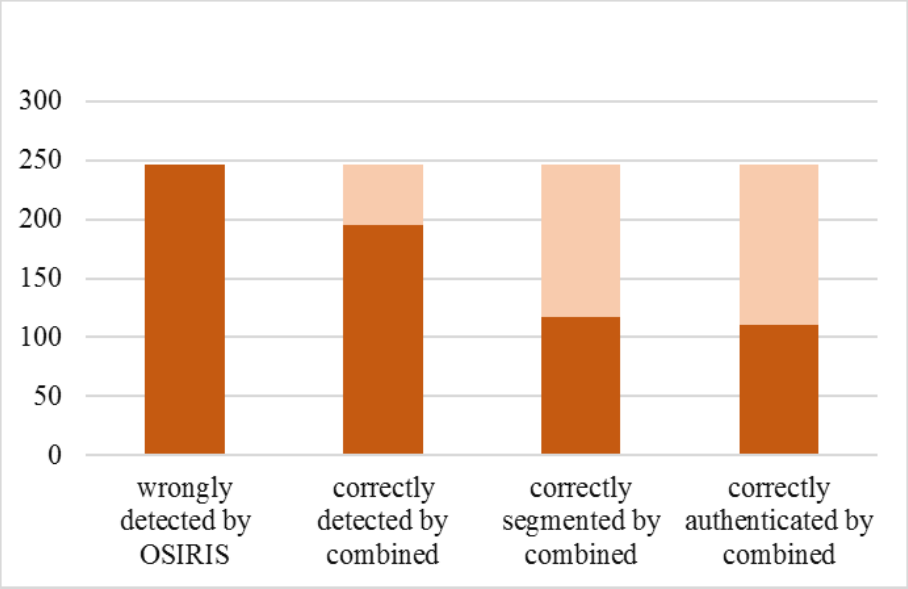


Figure 23. Iris authentication results for images wrongly detected by OSIRIS

Finally, all query images that are not in gallery images are checked for authentication results. Table 9 shows the authentication results for giving a full half face image and giving an eye region image cut by likelihood combination method to OSIRIS. Giving an eye region image as an input shows better performance than giving a half face input image. Therefore, eye region detection before segmentation stage is important to improve iris authentication. Our proposed likelihood combination method for eye region detection is an effective method for improving iris authentication.

Table 9. The overall iris authentication results

	Correctly Authenticated
OSIRIS	3068/3977(77.1%)
Combined	3191/3981(80.2%)

CHAPTER 6. CONCLUSION

In this research, we propose a likelihood combination method to detect an eye region from a face image. The proposed approach introduces two likelihoods, one for AKAZE feature matching and one for template matching, and these two likelihoods are combined. The maximum pixel value location is searched from the combined likelihood image and the eye region image is cut based on this highest location. Giving the cut eye image has greater possibility to correctly segment irises as it can avoid wrong segmentation on other organs from the face.

We apply our proposed method to CASIA-v4-Distance dataset and evaluate the performance. We firstly divide the face image into two left and right half face images. Then, the eye region is detected by likelihood combination method and cut the half image around this region. The cut image is given as input to the OSIRIS to proceed the iris authentication. The overall result shows that detecting eye region before applying OSIRIS can enhance the accuracy of iris authentication on face images acquired under less constrained environment.

In order to improve the authentication accuracy, we need to change the segmentation procedure. Although the likelihood combination method can detect eye region very well for images captured under less constrained environment, the segmentation procedure in this work cannot work well for images that have some noises such as specular reflection on the iris region. The improvement of iris segmentation method itself is left as a future work. Another future work is to compare with other eye detectors like Adaboost eye detector [31] proposed by Viola Jones to know how our proposed method is applicable and its strength and weakness.

REFERENCES

[1] Kitahara, Itaru, et al. "Thermal-ID: A Personal Identification Method Using Body Temperature". Adjunct Proceeding of the 3rd International Conference on Pervasive Computing(2005) : 69-72.

[2] Jain, Anil K., and S. Prabhakar. "Biometric authentication." Scholarpedia 3.6 (2008): 3716.

[3] Alcantarilla, Pablo F., and T. Solutions. "Fast explicit diffusion for accelerated features in nonlinear scale spaces." IEEE transactions on pattern analysis and machine intelligence 34.7 (2011): 1281-1298.

[4] Biometric Ideal test, CASIA-V4-Distance dataset, http://biometrics.idealtest.org/dbDetailForUser.do?id=4

[5] Faulds, Henry. "On the skin-furrows of the hand." Nature 22.574 (1880): 605.

[6] Li, S. Z. & Jain, A. K. eds., "Handbook of face recognition". Springer(2005).

[7] What are the danger of retina scanner, http://smallbusiness.chron.com/dangers-retinal-scanners-70631.html.

[8] Daugman, John. "How iris recognition works." IEEE Transactions on circuits and systems for video technology 14.1 (2004): 21-30.

[9] Iris image https://indranilsinharoy.com/2014/12/05/dissertation_series/.

[10] Garg, Surbhi, and Harmeet Kaur. "Survey Paper on Phase Based Iris Recognition." International Journal 4.4 (2014).

[11] Wildes, Richard P. "Iris recognition: an emerging biometric technology". Proceedings of the IEEE 85.9 (1997): 1348-1363.

[12] Masek, Libor, and Peter Kovesi. "Matlab source code for a biometric identification system based on iris patterns." The School of Computer Science and Software Engineering, The University of Western Australia 2.4 (2003).

[13] Lee, Yooyoung, Ross J. Micheals, and P. Jonathon Phillips. "Improvements in video-based automated system for iris recognition (vasir)". IEEE Workshop on Motion and Video Computing, WMVC'09 (2009).

[14] Othman, Nadia. "Fusion techniques for iris recognition in degraded sequences". Diss. Paris Saclay, (2016).

[15] Sutra, G., et al. "A biometric reference system for iris. OSIRIS version 4.1. Accessed: October 1, 2014." (2013), http://svnext.it-sudparis.eu/svnview2-eph/ref_syst//Iris_Osiris_v4.1/download/.

[16] Zhu, Zhiwei, Kikuo Fujimura, and Qiang Ji. "Real-time eye detection and tracking under various light conditions". The 2002 symposium on Eye tracking research & applications. ACM, (2002).

[17] Brunelli, Roberto, and Tomaso Poggio. “Face recognition: Features versus templates.” IEEE transactions on pattern analysis and machine intelligence 15.10 (1993): 1042-1052.

[18] Beymer, David J. “Face Recognition Under Varying Pose”. AI-M-1461. MASSACHUSETTS INST OF TECH CAMBRIDGE ARTIFICIAL INTELLIGENCE LAB, (1993) : Paper No. 89.

[19] Pentland, Alex, Baback Moghaddam, and Thad Starner. “View-based and modular eigenspaces for face recognition.” CVPR 1994(1994).

[20] Lam, Kin-Man, and Hong Yan. “Locating and extracting the eye in human face images”. Pattern recognition 29.5 (1996): 771-779.

[21] Yuille, Alan L., Peter W. Hallinan, and David S. Cohen. “Feature extraction from faces using deformable templates.” International journal of computer vision 8.2 (1992): 99-111.

[22] Feng, Guo Can, and Pong C. Yuen. “Multi-cues eye detection on grey intensity image”. Pattern recognition 34.5 (2001): 1033-1046.

[23] Z.H. Zhou, X. Geng, “Projection functions for eye detection”, Pattern Recognition 37 (5) (2004): 1049–1056.

[24] Kawaguchi, Tsuyoshi, and Mohamed Rizon, “Iris detection using intensity and edge information”, Pattern Recognition 36.2 (2003): 549-562.

[25] S.A. Sirohey, A. Rosenfeld, “Eye detection in a face image using linear and nonlinear filters,” Pattern Recognition 34 (2001): 1367–1391.

[26] J. Huang, H. Wechsler, “Visual routines for eye location using learning and evolution”, IEEE Trans. Evol. Comput. 4 (1) (2000): 73–82.

[27] J. Wu, Z.H. Zhou, “Efficient face candidates selector for face detection”, Pattern Recognition 36 (2003): 1175–1186.

[28] C.C. Han, H.Y.M. Liao, G.J. Yu, L.H. Chen, “Fast face detection via morphology-based pre-processing”, Pattern Recognition 33 (2000): 1701–1712.

[29] R.-L. Hsu, M. Abdel-Mottaleb, A.K. Jain, “Face detection in color images”, IEEE Trans. Pattern Anal. Mach. Intell. 24 (5) (2002): 696–706.

[30] Alcantarilla, Pablo, Adrien Bartoli, and Andrew Davison. “KAZE features”. ECCV 2012 (2012): 214-227.

[31] He, Zhaofeng, et al. “Toward accurate and fast iris segmentation for iris biometrics.” IEEE transactions on pattern analysis and machine intelligence, 31.9 (2009): 1670-1684.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest appreciation to my supervisor, Professor **Masayuki MUKUNOKI** for his patient guidance, valuable suggestions, enthusiastic encouragement and the continuous support throughout my research work. His precious instruction, guidance and eagerness on my thesis give strength to work out in all my research work. I am really appreciated for giving his time generously to me. Without him, this thesis would not possible. I also would like to thank Professor **Kunihito YAMAMORI** and **Dr. Zin Mar Kyu**, for their assistance on my thesis work.

I would like to give my sincere gratitude to the President of University of Miyazaki-Professor **Tsuyomu IKENOUE**, the members of my thesis committee, the professors for their applicable lectures that are appropriated for the research, the student affairs and GSO (International Relations Centre) family for their help to me.

I would like to express my immense sense of gratitude to Professor **Pyke Tin** and Professor **Thi Thi Zin** for their warmly carefulness like a family while I am studying in University of Miyazaki. I would like to say thanks to Myanmar DDP students and Myanmar Miyazaki Family for their valuable help to me whenever I need a help during staying in Miyazaki.

I also would like to appreciate to all our laboratory members for their warmest welcome, generous hospitality and sincerely friendship to me. And I would like to say deeply thanks to them for their kindness and giving time for all needed helps to me.

I would like to say thanks to CASIA (Chinese Academy of Science) Iris Dataset for their great help for my research work.

Furthermore, I would like to express my special thanks to all person in JASSO (Japan Student Service Organization) for their support to study in University of Miyazaki.

Lastly, I wish to appreciate to my beloved parents for their greatest kindness and endless love. I am heartily thankful to my parents, teachers and friends for their affection.

ABSTRACT

In this research, we propose a likelihood combination method to detect the eye region to improve iris authentication. For face images with some specular reflection, iris has more possibility to be wrongly segmented in the segmentation stage. In order to avoid wrong segmentation on other organs from the face, the accurate eye region is detected firstly by our proposed method before iris segmentation. The input face image is divided into two left and right half face images in order to detect the individual eyes. The likelihood images are created by AKAZE feature matching and template matching for each half input image. The feature points obtained from AKAZE feature matching are used to create the first likelihood image and the matched points from template matching are used to create the second likelihood image. Then, these two likelihood images are combined and the eye region is detected based on the highest peak from combined likelihood image. Giving a cut eye region image as input for the iris segmentation stage has more possibility to be correctly segmented by OSIRIS, a reference iris authentication system, as it contains iris area. Experimental results show that our likelihood combination method has good performance to detect eye region with 98.4% accuracy for CASIA-v4-Distance dataset and detected eye region images can authenticate better than giving face images as input.

Keywords - Eye region detection; AKAZE; Template matching; likelihood combination; OSIRIS

CONTENTS

PAGE

LIST OF FIGURES

LIST OF TABLES