An Accurate Iris Segmentation Framework under Relaxed Imaging Constraints using Total Variation Model

Zijing Zhao, Ajay Kumar
Department of Computing, The Hong Kong Polytechnic University
Hunghom, Kowloon, Hong Kong
jason.zhao@connect.polyu.hk, ajay.kumar@polyu.edu.hk

Abstract

This paper proposes a novel and more accurate iris segmentation framework to automatically segment iris region from the face images acquired with relaxed imaging under visible or near-infrared illumination, which provides strong feasibility for applications in surveillance, forensics and the search for missing children, etc. The proposed framework is built on a novel total-variation based formulation which uses l^1 norm regularization to robustly suppress noisy texture pixels for the accurate iris localization. A series of novel and robust post processing operations are introduced to more accurately localize the limbic boundaries. Our experimental results on three publicly available databases, i.e., FRGC, UBIRIS.v2 and *CASIA.v4-distance, achieve significant performance improvement in terms of iris segmentation accuracy over the state-of-the-art approaches in the literature. Besides, we have shown that using iris masks generated from the proposed approach helps to improve iris recognition performance as well. Unlike prior work, all the implementations in this paper are made publicly available to further advance research and applications in biometrics at-d-distance.*

1. Introduction

Iris recognition is one of the most accurate and widely employed approaches for the automated personal identification. The performance of iris recognition algorithms is highly dependent on the effectiveness of segmenting iris region pixels [7]. However, the traditional iris segmentation and feature matching approaches adopt only to near-infrared illumination and require the subjects to be sampled under strictly constrained condition [6], which is the major difficulty for deploying iris recognition system in civilian and surveillance applications on a larger scale. Automated iris segmentation has been a topic of considerable research in recent past [21]-[22] and many methods [2]-[6] have been proposed to address the problem. However the accuracy of currently available iris segmentation algorithms is still below the expectations and requires further improvement for the deployments.

This paper proposes a new framework to automatically and accurately segment iris images from the distantly acquired face images. The developed approach can robustly operate using face or eye images acquired under less-constrained environments, *i.e.*, using images acquired from a distance (typically 3-8m) and under near-infrared (NIR) or visible-wavelength (VW) illumination. The key contributions from this paper can be summarized as follows:

1. With the help of earlier studies on gradient dependent regularizer, such as relative total variation regularizer [10], we develop a new total variation formulation for iris segmentation in which the eye structure and surrounding texture are differently penalized. This formulation incorporates with an l^1 *norm* which is more effective and also computationally efficient. Our experimental results on three publicly available databases achieve significantly superiors results over previous approaches presented in the most recent literature [2]-[3]. Moreover, the method developed in this paper does not require any training and therefore is more attractive for the deployment in surveillance applications.
2. We develop a series robust post-processing operations to accurately localize limbic boundaries in noisy iris images. The adaptive and self-correcting methodology introduced in these operations can independently exploit the local features as much as possible, and helps to significantly reduce global errors. The post-processing operations can effectively use the intermediate results and adopt dynamic threshold mechanisms. Such robust strategies help to improve the overall accuracy in the segmentation of noisy iris images and can also be applied in other challenging problems in surveillances and remote sensing.

The performance of the proposed approach* have been evaluated on three publicly available databases, *i.e.*, UBIRIS.v2 [8], FRGC [17] using visible imaging and CASIA.v4-distance [16] under near infrared. The experimental results suggest average improvements of 28.82%, 30.98% and 16.05% on iris segmentation accuracy over state-of-the-art method on respective databases. Besides, we also illustrate from the experiments that using iris masks generated from our approach helps improve iris recognition performance.

1.1. Related Work

Most of the earlier work on iris segmentation uses NIR

* The implementation codes for our algorithm are available via [24].

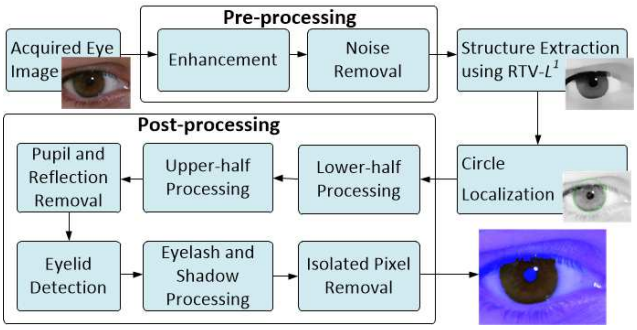
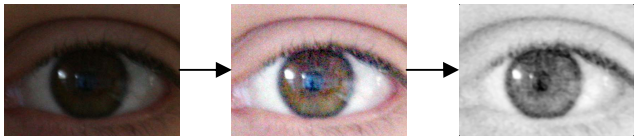


Figure 1: The block diagram for the proposed iris segmentation.

images which are acquired from close distances. Duagman's Integro-differential operator [1] is one of the most classical algorithms for iris segmentation under NIR illumination and is adopted in most of the commercial systems nowadays. It searches for a maximum response of an integro-differential expression and then locates the circle of iris. However, as explored and addressed in [6]-[7], *etc*., under VW illumination or less-constrained environment, quality of images drops and such traditional approach performs poorly.

The iris segmentation approach developed by Tan *et al.* [5] first adopts an iterative technique to cluster the iris and non-iris region coarsely, and then uses an improved integro-differential operator to locate the iris and pupil circle coarsely. One key limitation of this algorithm is that it relies highly on the coarse clustering result so that the final accuracy will be heavily affected if the first step is not accurate. Another promising approach by Proença [4] proposes to exploit local color features and classify iris pixels using a neural network. However, the color features are not very stable, which often leads to lower reliability. A recent work detailed in [3] also offers highly competitive alternative for the iris segmentation under less constrained imaging environment. This approach first adopts a Random Walker [13] to coarsely segment the iris region to locate the iris circle, then applies a set of gray level statistics based operations to refine the boundary. This method reports a better accuracy than previous ones. However, this approach also relies on the coarse segmentation result too much, and in its post-processing operations, one common threshold value is used for the whole iris, which may not fit local features and is possible to cause global error.

Another promising work in relevant domain has been proposed by Li and Savvides [2]. In this method, a Gaussian Mixture Model (GMM) was adopted to simulate iris pixel distribution and an unsupervised training method was used to obtain the parameters for the GMM. It has shown very high segmentation accuracy and reliability. However, a critical step for iris segmentation, which is the localization of iris and pupil circles, was performed manually in the experiments presented in this paper, while other methods mentioned above locate the circles automatically. In other words, the performance of [2] will highly depend on the accuracy of iris and pupil circle



(a) (b) (c)

Figure 2: Sample image from the pre-processing stage: (a) original image, (b) enhanced image, (c) smoothed red channel.

localization. In practice, iris and pupil circle localization is not only used in iris segmentation, but also necessary for the iris normalization, which unwraps the iris region into a polar coordinate system and is an essential step for most of the iris recognition algorithms.

2. Iris Segmentation Under Less-Constrained Imaging

This section details the methodologies used in the proposed iris segmentation approach. The overall framework of the developed approach is illustrated in Figure 1. The proposed approach adopts a coarse-to-fine strategy to segment iris region pixels from the background (region pixels surrounding the iris) and foreground (noisy pixels in the iris region) pixels in the acquired eye images. Our approach assumes that each of the eye images may be acquired under a relaxed imaging environment, *i.e*., at-a-distance and under variable spectrum bands.

2.1. Preprocessing

Under less-constrained imaging, several factors such as varying illumination intensity and the angle of the illumination source can have adverse impact on the accuracy and quality of iris segmentation. Such unexpected changes yield severe challenges in not only the iris biometrics but also many other image understanding tasks. We use the Single Scale Retinex (SSR) approach [12] for normalizing eye image illumination. The SSR enhancement method is able to improve color consistency under severe illumination variance. A sample image after applying SSR enhancement is shown in Figure 2 (b).

After enhancement, we apply a median filter on the image to suppress isolated noisy pixels. Moreover, we only use the red channel in the following process because the imaging spectrum of red channel is closest to NIR, which retains better image quality. In Figure 2 (c) we can see a sample result from the pre-processing stage.

2.2. Total Variation-Based Iris Structure Extraction

One common characteristic for the eye images acquired under less-constrained environments is the sensitivity to noisy and complex details such as reflection and eyelashes, which are not needed in the initial structure analysis. The above factors are the major reason why the traditionally effective integro-differential operator or circular Hough

transform perform poorly on images acquired under less-constrained environments, because both methods require clear contrast of structure components and least interference from noise. We exploit the total variation (TV) model to address such a problem. There have also been studies on using the total variation model for other biometric segmentation problems such as fingerprint segmentation [18].

2.2.1 Theoretical Foundation of Total Variation Model

There are several total variation (TV) regularizers for image structure separation in the literature, of which most are extended from TV-L^2 [14]. A recent reference in [10] proposed relative total variation (RTV) to measure and regularize local pixel variation. Such local gradient descriptors offer the strong capability to distinguish key image structure from the background image details. Motivated by such prior studies, we propose to use an improved RTV model to first localize the key eye structure, *i.e.*, eyelid, pupil and sclera boundaries, in the noisy eye images. Such localization of eye structure can be used to accurately locate pupillary and limbic boundaries for accurate iris segmentation. In the following, we provide brief review on the theoretical principles of RTV which are later used to develop an improved RTV model incorporated with l^1 norm regularization to more effectively locate eye structure of key interest.

The windowed total variation of an image S within a local rectangle region R is expressed as follows:

$$\begin{aligned} D_{S,x} &= G_\sigma * |\partial_x S| \\ D_{S,y} &= G_\sigma * |\partial_y S| \end{aligned} \tag{1}$$

where G_σ is a Gaussian kernel with standard deviation σ, ∂_x and ∂_y are the partial derivatives on image S in two directions and $*$ represents the convolution operation. By the convolution, which gives a weighting sum of nearby absolute gradients, we can observe that $D_{S,x}$ and $D_{S,y}$ represent absolute spatial difference within a rectangular window. In earlier studies in [10], both the detail and structure patches in an image with salient textures yield large D, which indicates that the windowed total variation is responsive to visual saliency.

Another effective measure to help distinguishing prominent structures from the texture elements is to use windowed inherent variation, expressed as:

$$\begin{aligned} L_{S,x} &= |G_\sigma * \partial_x S| \\ L_{S,y} &= |G_\sigma * \partial_y S| \end{aligned} \tag{2}$$

Different from D, L measures overall spatial variation because $\partial_x S$ and $\partial_y S$ may be positive or negative, and therefore such values may eliminate or offset others by the convolution in frequently varying gradient region. As a result, structure patches are typically expected to yield larger L than those from texture patches.

The contrast between texture and structure can be further enhanced by combining D and L as RTV, expressed as follows:

$$RTV_{S,p} = \frac{D_{S,x}(p)}{L_{S,x}(p)+\varepsilon} + \frac{D_{S,y}(p)}{L_{S,y}(p)+\varepsilon} \tag{3}$$

where p is the pixel index, ε is a small positive number to avoid division by zero. From expression (3) we can observe that texture region is typically expected to yield larger RTV than structure since the denominator of the formulation, L, responses smaller value for texture. Making use of such a property of RTV, reference [10] proposed to minimize following energy to remove the texture (*e.g.*, details and noise) from the input image:

$$\arg\min_S \sum_p \lambda \cdot RTV_{S,p} + \left(S_p - I_p\right)^2 \tag{4}$$

where I is the input image and S is the output image. Notice that equation (4) incorporates the square of an l^2 norm to enforce the similarity between the input and output image, which is similar to many other variants of TV regularization. We will refer to such a method as RTV-L^2 for short.

2.2.2 Extracting Eye Structure Using RTV-L^1

Each of the iris images acquired for conventional iris recognition includes surrounding eye structure. This structure essentially includes curved regions representing eyelid, pupil and sclera boundaries. Our objective is to locate the iris by automatically extracting such elements representing eye structure and other non-structural elements such as eyelash, and iris texture can be treated as noise because they could have interference on our iris localization. Therefore, the RTV-L^2 approach which can remove details and texture while maintaining main structure of the input image is a good choice for our purpose. However, it has been studied in several references [20], [21] that using l^1 norm instead of l^2 in such energy regularizers has better performance in some applications and presents more important geometric properties. We have studied the difference between l^1 and l^2 norm in RTV regularization, and propose to adopt l^1 norm instead of the original l^2 norm, i.e., we solve the following problem which we refer to as RTV-L^1:

$$\arg\min_S \sum_p \lambda \cdot RTV_{S,p} + \left|S_p - I_p\right| \tag{5}$$

The difference between the output images by solving problems (4) and (5) is illustrated in Figure 3. We can observe from Figure 3 that while both RTV-L^1 and RTV-L^2 can suppress texture and noise, the results from RTV-L^1 are sharper at critical edges than those from RTV-L^2. This confirms the arguments that using l^1 norm in the energy regularizer can present more important geometric properties, which is considered helpful for the subsequent iris localization process. The detailed numeric solution for problem (5) will be introduced in following sections.

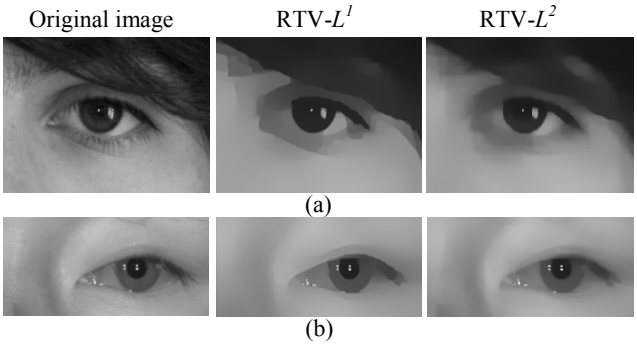


Figure 3: Sample results of RTV-L^1 and RTV-L^2 for eye images under (a) visible illumination and (b) NIR illumination.

2.2.3 Numeric Solution for RTV-L^1

The objective function in problem (5) is non-linear and non-convex. A trivial solution for this problem is not available. In addition, by replacing the l^2 norm with l^1 norm, the structure of the objective function has changed that the approximating solution proposed in [10] becomes unusable. Here we propose an effective dual formulation based solution similar to [19] for the RTV-L^1 problem. First, we approximate the minimization for problem (5) as minimizing the following new problem:

$$\arg\min_{S,V} \sum_p \lambda \cdot RTV_{S,p} + \frac{1}{2\theta}\left(S_p + V_p - I_p\right)^2 + \left|V_p\right| \quad (6)$$

where V is a new variable in matrix form and the positive parameter θ is small, thus we have $V \approx I - S$. As a result, S presents the structural information and V captures the texture information from the input image. The minimization for problem (6) is performed with respect to S and V separately and iteratively. Thus, it boils down to the following two sub-problems:

(i) S being fixed, search for V for the problem:

$$\arg\min_{V} \sum_p \frac{1}{2\theta}\left(S_p + V_p - I_p\right)^2 + \left|V_p\right| \quad (7)$$

(ii) V being fixed, search for S for the problem:

$$\arg\min_{S} \sum_p \lambda \cdot RTV_{S,p} + \frac{1}{2\theta}\left(S_p + V_p - I_p\right)^2 \quad (8)$$

Problem (7) and (8) are solved alternately and iteratively, and then the energy function in problem (6) keeps reducing until it converges to a satisfying level. Following we will give solutions for (7) and (8):

(a) Solution for (7):

Since the objective function at each pixel is independent from others, this problem is a 1-D minimization problem and can be easily solved by calculus. The solution is given by:

$$V_p = \begin{cases} I_p - S_p - \theta & \text{if } I_p - S_p > \theta \\ I_p - S_p + \theta & \text{if } I_p - S_p < -\theta \\ 0 & \text{if } |I_p - S_p| \le \theta \end{cases} \quad (9)$$

Such solution is also given in [19].

(b) Solution for (8):

The objective function in problem (8) has a quadratic term, which is very similar to the original RTV-L^2 problem in [10]. Therefore, we can use a similar iterative solution proposed in [10] to solve problem (8) approximately. As shown in [10], the objective function in (8) can be approximated with a matrix form:

$$\begin{aligned} &\sum_p \lambda \cdot RTV_{S,p} + \frac{1}{2\theta}\left(S_p + V_p - I_p\right)^2 \approx \\ &\left(v_S - v_{I-V}\right)^T \left(v_S - v_{I-V}\right) + \\ &2\theta\lambda \cdot {v_S}^T \left({C_x}^T U_{S,x} W_{S,x} C_x + {C_y}^T U_{S,y} W_{S,y} C_y\right) v_S \end{aligned} \quad (10)$$

where v_Q is the vector representation of matrix Q, $C_{x(y)}$ is a Toeplitz matrix from gradient operator in x or y direction. $U_{S,x}$ and $W_{S,x}$ are diagonal matrices, whose values on the diagonals are respectively

$$\begin{aligned} U_{S,x}[p,p] &= (G_\sigma * \frac{1}{\left|G_\sigma * \partial_x S\right| + \varepsilon})_p \\ W_{S,x}[p,p] &= \frac{1}{\left|(\partial_x S)_p\right| + \varepsilon'} \end{aligned} \quad (11)$$

where p is the pixel index in the vector representation of the image, ε and ε' are newly introduced small positive constants for preventing division by zero. After the approximation, let:

$$L = {C_x}^T U_{S,x} W_{S,x} C_x + {C_y}^T U_{S,y} W_{S,y} C_y \quad (12)$$

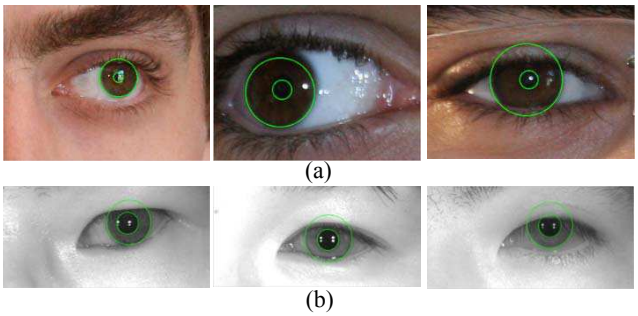
Considering L as a constant and compute the value of L using the results from last iteration, then the minimization problem (8) boils down to the following:

$$(\mathbf{1} + 2\theta\lambda L) \cdot v_S = v_{I-V} \quad (13)$$

The problem in (12) is easy to be solved using knowledge of linear algebra. As the number of iteration increases, the output approaches to the optimal solution and the value of the energy function in (6) keeps reducing until it converges to a stable level. Currently we iterate five times for each eye image based on the observation on the output and receive satisfying noise removal effect, as shown in Figure 3.

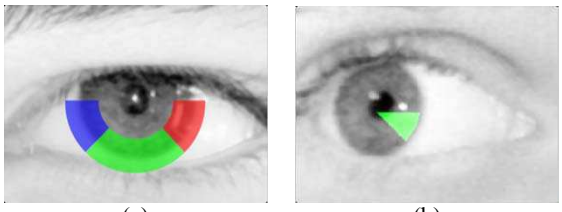
2.3. Coarse Iris Localization Using a Circle

As discussed in section 1.1, a simple circular model cannot be employed to accurately segment iris images acquired under less constrained environments. However, it is widely observed that the human iris can be coarsely approximated as a circle [1], [15]. A circular boundary that coarsely but closely fits the limbic boundary can be used to further refine the boundaries for accurate iris segmentation using a series of efficient post-processing algorithms. In this paper, we refer to such a coarse localization circle as an *iris circle*. Similarly, the *pupil circle* describes the circular boundaries that coarsely fit the pupillary boundary of iris images acquired for the segmentation.



(a) (b)

Figure 4: Sample results from the iris and pupil circle localization for (a) VW images and (b) NIR images.



(a) (b)

Figure 5: Illustration of three sectorial regions to be processed (a) and the Otsu's thresholding result for one sectorial region (b).

After structure extraction, the noise of the eye images is highly suppressed and it is possible to use the circular Hough transform (CHT) based approach to detect the iris and pupil circles coarsely, which highly relies on the clarity of the image structure. We implemented an improved version of CHT based on the two-phase CHT introduced in [9]. Firstly, we only detect the lower half circles to prevent possible interference from the eyelashes or eyebrow. Secondly, after the first phase in [9] which estimates the circle center, we enabled re-searching for the circle center within a rectangular region around the estimated center, to more accurately detect the center position and the radius. The robustness for coarsely localizing the iris region increases with the improved CHT. We detect the circles with empirically proper radius ranges, whose sample results are shown in Figure 4. The possible ranges of radius for the databases we used, *i.e.*, UBIRIS.v2, FRGC and CASIA.v4-distance, are [35, 120], [25, 40] and [60, 100] respectively.

2.4. Iris Pixel Identification by Local Gray Level Analysis

Automated boundary refinement approach has to be developed to accurately identify the limbic boundaries after the *iris circle* is detected. We developed an adaptive histogram-based binarization approach to firstly process lower half pixels of the *iris circle* in the image.

2.4.1 Adaptive Detection of Lower Half Iris and Sclera Boundary

The reason for processing lower part firstly is that the lower half iris is less likely to be affected by eyelash and eyelid. Therefore, accurately identifying the iris pixels in the lower half region is firstly considered in identifying noisy pixels

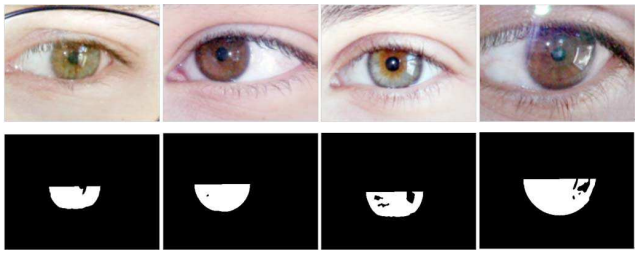


Figure 6: Sample iris images and corresponding results from post processing of lower half iris pixel region.

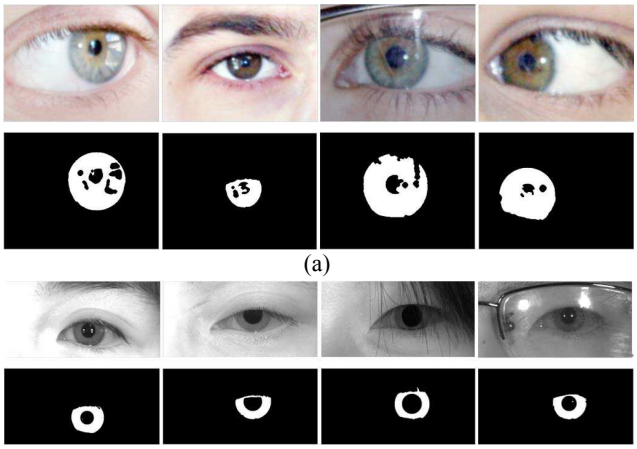
using the thresholding. Firstly processing the lower half not only can improve segmentation accuracy but also help to detect the thresholds for accurately segmenting the upper half.

The lower half circle is firstly processed by performing N sector thresholdings. In one thresholding, pixels in a certain sectorial region as expressed in the following are identified:

$$C_{\phi_1,\phi_2} = \left\{ p \;\middle|\; t_1 r_{ir} \le \left|\overrightarrow{cp}\right| \le t_2 r_{ir} \text{ and } \phi_1 \le \theta_p \le \phi_2 \right\} \quad (14)$$

where c and r_{ir} are the center and radius of *iris circle* respectively, θ_p is the angle from *x axis* to the vector $\overrightarrow{cp}$, says the central angle at point p, $[\phi_1, \phi_2]$ is the range of central angles with $0 \le \phi_1 < \phi_2 \le \pi$, $[t_1, t_2]$ is the constant ratio range to the iris radius restricting the region of the sector, and is empirically set to [0.6, 1.35]. In our approach, N is set to 3, and the sequence of ranges of central angels are $[0, \frac{\pi}{4}]$, $[\frac{\pi}{4}, \frac{3\pi}{4}]$ and $[\frac{3\pi}{4}, \pi]$ respectively. These sectorial regions are also shown in Figure 5 (a).

If the edge is clear and the *iris circle* is accurate, we can choose a threshold value that separates the low end and high end of the pixel values inside the sectorial region. Otsu's method is a good approach for such purpose. It can automatically locate valley point between two peaks in the histogram of a set of pixel values using two-class separation metric. The significant aspect of our strategy is that we adopt different threshold values at each of the different sectors, which ensures that the overall error in the identification of iris pixels is significantly reduced. In addition, the number of sectors and range of angle sequences can be varied to accommodate iris images of degraded quality. Note that the acquired eye images suffer from serious noise and occlusions. If each segment is too small (N is large), the computed threshold may not be robust and the computational time will also increase. Therefore, $N = 3$ is a reasonable choice. In [3], the threshold is obtained from statistical information of pixel values within a region near the pupil. Such method applies only one fixed threshold for the whole circular boundary, which may not fit local features very well. Figure 6 shows the sample results from the post-processing of lower half of iris region pixels.



(a)

(b)

Figure 7: Sample results after upper half masking, pupil removal and source reflection removal for (a) VW images and (b) NIR images.

2.4.2 Coarse-to-Fine Localization for Upper-Half Iris, Pupil Region and Reflection

The upper half part is expected to be highly noisy, which is caused by the eyelash and shadow, and quite a significant part of the iris is occluded by eyelid and therefore the sector thresholding may not work well here. We can reuse the previous thresholds from the sectorial thresholding described in section 2.4.1. We just segment the upper-left 1/4 circle using threshold determined in $C_{\frac{3\pi}{4},\pi}$, and the upper right one using threshold determined in $C_{0,\frac{\pi}{4}}$. This approach is not expected to cause big error because two pairs of these regions are continuously connected, and further refinement regarding eyelid, eyelash and shadow will be performed.

Since we have already detected the *pupil circle* earlier, the pupil removal step is to eliminate the pupil region pixels from the iris circle in previous step. Another effect of the sector thresholding is that the detected threshold can be used to identify source reflections that usually exist in the images acquired under less-constrained imaging environment and occlude the iris region. We eliminate pixels whose gray levels are higher than the highest threshold among all the three sectors in the lower half iris processing section (Figure 5). In summary, the pixels which are brighter than the brightest pixels in lower half of iris region are considered as source reflection. Figure 7 illustrates some sample results after masking upper half iris, eliminating pixels belonging to pupil and source reflection.

2.4.3 Identifying Eyelid, Eyelash and Shadow (ES)

As discussed earlier, the ES region brings much noise and ambiguity in the segmentation process. It is important to carefully identify this restricted region to perform any refinement. Therefore, the position of upper eyelid should be accurately located.

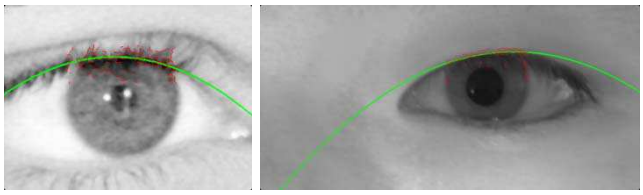


Figure 8: Sample results of the proposed eyelid fitting approach. Green curve is the fitted parabola representing upper eyelid, and the red points are the edge points detected by the canny edge detector in region *R*.

1. Eyelid Fitting

Using a parabola to approximate the eyelids is a popular approach in many iris segmentation algorithms and is found to have higher performance than other approaches [26]. Therefore we also propose to fit the eyelid with a parabola, which is in the following form:

$$y - c = a(x - b)^2 \tag{15}$$

Considering the shape of the upper eyelid and in order to fasten the parameter searching, we limit the ranges of *a*, *b* and *c* as follows:

$$\begin{cases} 0 < a < 1/r_{ir} \\ x_c - 2 \cdot r_{ir} < b < x_c + 2 \cdot r_{ir} \\ y_c - 1.5 \cdot r_{ir} < c < y_c - 0.3 \cdot r_{ir} \end{cases} \tag{16}$$

where (x_c, y_c) and r_{ir} are the center and radius of *iris circle* respectively. The range of a ensures that the parabola is orienting downwards and will not be too sharp, and the ranges of b and c make the vertex of the parabola not too far away from the iris.

The approach we propose to search the parabola is simple and yet very effective in terms of speed and accuracy. First, we define a rectangular region as the candidate eyelid area as follows:

$$R = \{(x, y) \,|\, x_c - r_{ir} \le x \le x + r_{ir}, y_c - r_{ir} \le y \le y_c - 0.3 \cdot r_{ir}\} \tag{17}$$

A canny edge detector is applied in R and let us donate the set of detected edge points as *E*. We assume that among the edge points in *E*, some are close to the position of eyelid, which we refer to as eyelid points, and some points belong to noise such as eyelashes and shadow, which we refer to as non-eyelid points. The spatial distribution of the non-eyelid points is highly random and less regular, while the positions of the eyelid points are very close to the parabola that can accurately fit the real eyelid. Therefore, we search for a parabola with the parameters $\{a,b,c\}$ that has maximum number of points in *E* lying on it. Moreover, we actually search for $\{a,b,c\}$ at discrete interval so the speed can be greatly fastened. Figure 8 illustrates two sample results of the proposed eyelid fitting approach, which is highly accurate.

2. Eyelash and Shadow Processing

Having located the upper eyelid, the next step is to mask

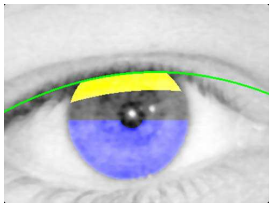


Figure 9: Illustration of ES processing. Pixels in the blue region are collected to calculate thresholds to process the pixels in the yellow region.

Table 1: Summary of databases employed in the experiments.

	UBIRIS.v2	CASIA.v4-distance	FRGC
Imaging illumination	visible	near-infrared	visible
Standoff distance	4 – 8m	⩾3m	N/A
Eye image size	400×300	about 780×400	300×150
No. of subjects	171	77	163
No. of images	1,000	581	540

out those pixels which are belonging to the eyelashes and shadow at a certain distance below the eyelid. This step is the same as described in [3] and we choose the distance as $0.3\times r_{ir}$. The pixel values within lower half of the currently processed iris mask are used to detect thresholds to identify those belonging to ES region. Figure 9 illustrates the idea.

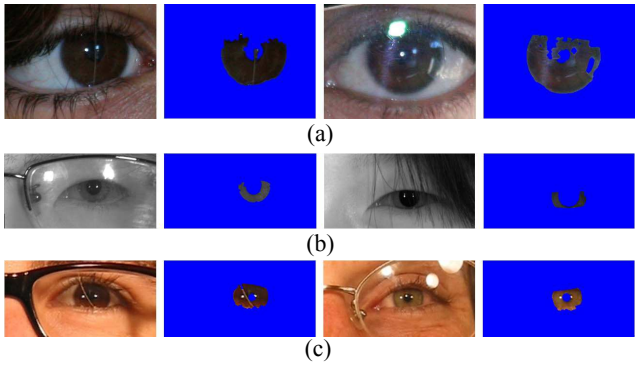
We choose the limiting thresholds that exclude 1% of the darkest pixels and 20% of the brightest pixels as the low and high thresholds respectively. Only the pixels between these two thresholds are retained. In order to eliminate isolated noisy pixels, the iris mask is subjected to an opening operation. Figure 10 illustrates some sample iris segmentation results from the databases used in this work.

3. Experiments and Results

3.1. Databases

We have used three publicly available databases, UBIRIS.v2 [8], FRGC [17] and CASIA.v4 [16] to perform the experiment for the iris segmentation and recognition under VW and NIR imaging. The images from these databases were acquired under less-constrained environment. It is judicious to expect that good performance on these databases indicates higher probability for the proposed approach to work well in surveillance and forensics applications. The summary of the employed subsets is presented in Table 1. We selected these subsets subject to the availablity of the ground truth iris masks (explained in section 3.2). Please see enclosed *supplementary file* for more details on eye detection and the experimental process.

As for the parameter tunning, there are mainly two types of parameters. The first one is those related to the proposed RTV-L^1 solution. We use the same set of parameters ($\lambda=0.2,\theta=0.05,\sigma=3,\varepsilon=\varepsilon'=0.005$) for all three databases, which illustrates that the proposed RTV-L^1 is highly generalizable for images captured in various



(a) (b) (c)

Figure 10: Sample source images and corresponding final segmentation results (non-iris region is masked with blue color) for (a) VW images from UBIRIS.v2, (b) NIR images from CASIA.v4-distance and (c) VW images from FRGC.

Table 2: Comparison of average segmentation error rates for different approaches.

	Iris Segmentation Error, $\bar{e}$ (%)		
Approaches	**UBIRIS.v2**	**CASIA.v4 -diatance**	**FRGC**
Proposed RTV-L^1	1.21	0.68	1.27
RTV-L^2	1.41	0.75	1.28
T-PAMI, 2013, [2]	1.92	0.85	1.34
T-IP, 2013, [3]	1.70	0.81	1.84
T-IP, 2012, [6]	1.90	1.13	1.84
T-PAMI, 2009, [4]	3.75	1.61	2.42
ImVis, 2010, [5]	3.49	1.71	3.30

condition. Other parameters are mainly database-specific, such as the range of radius of iris circle. Such parameters should be adjusted according to the image resolution.

3.2. Performance Evaluation

3.2.1 Segmentation Accuracy

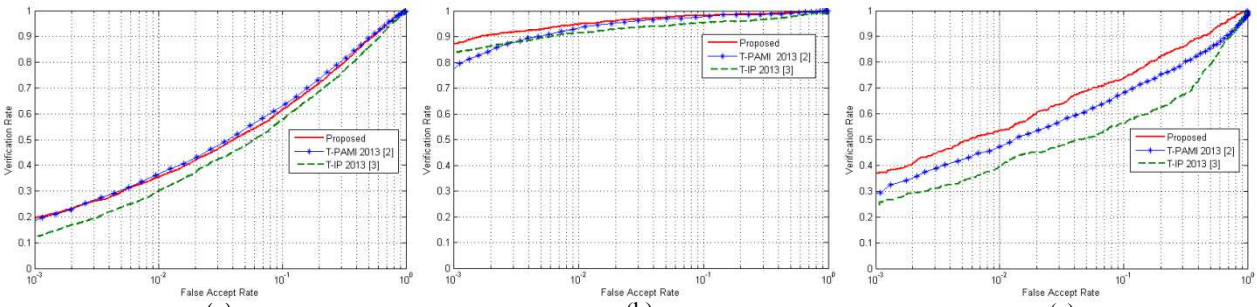
The accuracy of iris segmentation is evaluated using the same protocol as in the NICE. I competition [11], in which the average segmentation error rate is computed as follows:

$$\bar{e}=\frac{1}{N\times w\times h}\sum_{x\in w}\sum_{y\in h}T(x,y)\oplus M(x,y) \qquad (18)$$

where N is the total number of images, w and h are width and height of one image, T and M are the ground truth mask and generated iris mask respectively. The symbol $\oplus$ represents an exclusive OR operation to identify the segmentation error. While ground truth of UBIRIS.v2 are manually labeled and publicly provided by NICE.I, ground truth for the other two datasets is also manually generated by authors of [3] and made publicly available. Therefore, we can use the NICE.I protocol for the consistent segmentation accuracy evaluation.

Table 2 provides summary of the performance from state-of-the-art approaches in the recent literature while using above protocol†. The proposed approach achieved

† The average error rate of algorithm in [2] is also produced from our implementation and is made available via [24].



(a) (b) (c)

Figure 11: ROC curves of iris recognition experiments using iris masks generated from different segmentation approaches for (a) UBIRIS.v2, (b) CASIA.v4-distance and (c) FRGC.

average segmentation error rates of 1.21%, 0.70% and 1.29% for UBIRIS.v2, CASIA.v4-distance and FRGC respectively.

It can be observed from the statistics that the proposed approach consistently outperforms other iris segmentation methods developed in the literature. As compared with the recent approach published in [3], the proposed method can achieve average improvement of 28.82%, 16.05% and 30.98% for UBIRIS.v2, CASIA.v4-distance and FRGC databases respectively, in iris segmentation accuracy. It may be noted that the method described in [5] was *ranked first* in NICE I competition [11] and therefore provides a good benchmark for the comparison. We have also evaluated the performance when using the original RTV-L^2 approach for structure extraction and keeping other steps exactly the same. The results in Table 2 show that the proposed RTV-L^1 has noticable superiority over RTV-L^2 due to its ability to preserve sharpness of important edges.

3.2.2 Recognition Performance

The recognition performance is always the first concern in iris recognition systems. In order to ascertain that our accurate iris segmentation approach can also contribute to improving recognition performance, we have performed the experiments on iris recognition as well. We adopt the 1D log-Gabor filter as the feature encoding method, which is widely used in the deployed iris recognition systems, and use iris masks generated from different segmentation approaches for comparison. The parameters of the log-Gabor filter are optimized from a separate training set and kept the same within one database for different masks. Therefore the only factor that impacts the recognition performance is the iris segmentation approach. The training and testing protocols are detailed in *supplementary file* [24].

The Receiver Operating Characteristic (ROC) curves for the employed datasets using iris masks from comparative approaches are shown in Figure 11. From Figure 11 we can see that the experiments using the proposed iris segmentation approach produce better ROC than those using other segmentation approaches, clearly for FRGC and CASIA.v4-distance. For UBIRIS.v2, the proposed approach also improves the verification rate at lower false accept rate (FAR). Above experiments illustrate that the proposed iris segmentation approach not only provides the best segmentation accuracy but also offers noticeable improvements in the final iris recognition performance.

4. Conclusions

This paper has developed a more accurate iris segmentation framework to automatically segment iris image acquired under less-constrained imaging environment. The proposed approach introduces a new total-variation based energy regularizer incorporated with an l^1 norm, in which the slowly varying components of image structure such as eyelid, limbic boundaries, etc., and the surrounding texture and noise are differently penalized. In addition, an efficient solution for the proposed energy regularizing formulation is given. Such an approach allows us to reliably extract the eye structure for more accurately localizing iris and pupil circles for further segmentation. Our work also introduced a series of novel post-processing operations that exploit local (but often varying) distribution characteristics to adaptively refine pupillary and limbic boundaries. The overall framework has shown to be highly robust to achieve significant improvement in segmentation accuracy as well as iris recognition performance from publicly available iris databases that are under both VW and NIR spectrum.

The RTV-L^1 texture removal approach introduced in this paper is not only significant for the noisy iris segmentation but can also be potentially employed to solve other texture or object segmentation tasks which require removal of accompanying noise. The adaptive local intensity analysis developed in our work has been greatly successful in increasing the robustness of the proposed approach under less-constrained imaging. Such adaptive decision-making strategies can also be effectively used in other challenging problems in surveillances and remote sensing that often suffer from less stable illumination conditions and unwanted occlusions. The framework developed in this work provides robust and effective prerequisite for researchers and applications which attempt to perform accurate iris recognition on noisy images acquired under less-constrained environment and at-a-distance.

Acknowledgment

This work is supported by the General Research Fund from the Hong Kong Research Grant Council grant no. 15206814, (PolyU 152068/14E).

References

[1] J. Daugman, "How iris recognition works," *IEEE Trans. Circuits Syst. Video Technol*., vol. 14, no. 1, pp. 21–30, 2004.

[2] Y. H. Li and M. Savvides, "An automatic iris occlusion estimation method based on high-dimensional density estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 4, pp. 784-796, 2013.

[3] C. W. Tan and A. Kumar, "Towards online iris and periocular recognition under relaxed imaging constraints," *IEEE Trans. Image Process.,* vol. 22, no. 10, pp. 3751-3765, 2013.

[4] H, Proença, "Iris recognition: On the segmentation of degraded images acquired in the visible wavelength," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1502-1516, 2010.

[5] T. Tan, Z. He and Z. Sun, "Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition," *Image Vision Computing*, vol. 28, no. 2, pp. 223–230, 2010.

[6] C. W. Tan and A. Kumar, "Unified framework for automated iris segmentation using distantly acquired face images," *IEEE Trans. Image Process*., vol. 21, no. 9, pp. 4068-4079, 2012.

[7] J. K. Pillai, V. M. Patel, R. Chellappa and N. K. Ratha, "Secure and robust iris recognition using random projections and sparse representations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 9, pp. 1877-1893, 2011.

[8] H. Proenca, S. Filipe, R. Santos, J. Oliveira, and L. A. Alexandre, "The ubiris. v2: A database of visible wavelength iris images captured on-the-move and at-a-distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1529-1535, 2010.

[9] E. R. Davies. "Circle and ellipse detection" in *Computer and Machine Vision, Fourth Edition: Theory, Algorithms, Practicalities*, Waltham: Academic Press, 2012.

[10] L. Xu, Q. Yan, Y. Xia and J. Jia, "Structure extraction from texture via relative total variation," *ACM Trans. Graphics*, vol. 31, no. 6 (139), 2012.

[11] NICE.I - Noisy Iris Challenge Evaluation, Part I. `http://nice1.di.ubi.pt/`

[12] D. H. Brainard and B. A. Wandell, "Analysis of the retinex theory of color vision," *J. Optical Soc. Am. A*, vol. 3, no. 10, pp. 1651-1661, 1986.

[13] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768-1783, 2006.

[14] L. I. Rudin, S Osher and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D,* vol. 60, no. 1, pp. 259-268, 1992.

[15] A. Kumar and A. Passi, "Comparison and combination of iris matchers for reliable personal authentication," *Pattern Recognition*, vol. 43, no. 3, pp. 1016-1026, 2010.

[16] [38] Biometrics Ideal Test, CASIA.v4 database: http://www.idealtest.org/dbDetailForUser.do?id=4

[17] Face Recognition Grand Challenge, FRGC database: `http://www.nist.gov/itl/iad/ig/frgc.cfm`

[18] J. Zhang, R. Lai and C. C. Kuo,"Adaptive directional total-variation model for latent fingerprint segmentation," *IEEE Trans. Info. Forensics & Security*, vol. 8, pp. 1261-1273, 2013.

[19] X. Bresson, S. Esedoglu, P Vandergheynst, J. P. Thiran and S. Osher, "Fast global minimization of the active contour/snake model," *J. Math. Imaging & Vision*, vol. 28, no. 2, pp. 151-167, 2007.

[20] S. Alliney, "A property of the minimum vectors of a regularizing functional defined by means of the absolute norm," *IEEE Trans. Sig. Process*., vol. 45, no. 4, pp. 913-917, 1997.

[21] A. Kumar, T.-S. Chan, "Iris recognition using quaternionic sparse orientation code (QSOC)," *Proc. CVPR 2012*, pp. 59-64, CVPRW 2012, Providence, June 2012.

[22] A. Kumar, T.-S. Chan, C. W. Tan, "Human identification from at-a-distance face images using sparse representation of local iris features," *Proc. ICB 2012*, pp. 303-309, Apr. 2012.

[23] M. J. Burge and K. W. Bowyer. *Handbook of Iris Recognition*. Springer, 2013.

[24] Weblink to download implementation codes for this paper, http://www.comp.polyu.edu.hk/~csajaykr/tvmiris.htm, 2015

[25] M. Nokolova, "A variational approach to remove outliers and impulse noise," *J. Math. Imaging & Vision*, vol. 20, pp. 99-120, 2004.

[26] T. H. Min and R. H. Park, "Comparison of eyelid and eyelash detection algorithms for performance improvement of iris recognition," *Proc. ICIP*, pp. 257-260, San Diego, Oct. 2008.