Adaptive and Localized Iris Weight Map for Accurate Iris Recognition under Less Constrained Environments

Chun-Wei Tan, Ajay Kumar
Department of Computing, The Hong Kong Polytechnic University, Hong Kong
Email: cscwatan@comp.polyu.edu.hk, ajaykr@ieee.org

Abstract
Accurate iris recognition from the distantly acquired face or eye images requires development of effective strategies which can account for significant variations in the segmented iris image quality. Unlike conventional stop-and-stare mode iris recognition, iris images acquired under less constrained imaging environment specially those under visible illumination, are degraded by multiple sources of noise. Such accompanying noise is often embedded in the feature space (say iris code), during the feature extraction phase following the image normalization, and is highly correlated with the consistency of resulting iris bits. The bit consistency can be learned from the temporal iris codes such that the stable bits are considered less likely to be corrupted by noise, or vice versa. The relationship between the bit consistency and the accompanying noise is proposed to be modeled using a non-linear relationship that can follow a power-law. Therefore the noise perturbed bits are given less emphasis while consistent bits are given higher weight. The superiority of proposed iris matching strategy is ascertained by providing comparison with other state-of-the-art algorithms using publicly available databases: CASIA.v4-distance and UBIRIS.v2. Our experimental results suggest improvement of 13.9% and 47.2% in the average rank-one recognition respectively for the CASIA.v4-distance and UBIRIS.v2 databases.

1. Introduction
Iris recognition has emerged as one of the most promising technologies to provide reliable human identification. The existing state-of-the-art iris recognition algorithms have reported remarkable recognition accuracy on the iris images acquired using near infrared (NIR) imaging from controlled environment [1]-[4]. Such superiority has made iris recognition as one of the good candidates for large-scale human identification applications [5]. One major endeavor in the recent development of iris recognition technologies is to break through the practicality limitations of the existing iris recognition systems to allow the image acquisition at-a-distance and from less controlled conditions [5], [10], [18]. Such properties are essentially desired especially to meet the increasingly demand for forensic and high security surveillance applications [7]-[8], for e.g. provide critical early warning to thwart terrorism or other human-based threats. Along with this initiative, several iris databases composed of iris images acquired under less constrained environments have been released in public domain in order to encourage further research in this area. Among such databases, e.g. [18], there have been attempts which employed visible imaging for image acquisition in order to overcome several limitations of the existing commercial NIR-based iris recognition systems. Most importantly, the existing NIR-based iris recognition approaches require high degree of active cooperation from the subjects to provide their iris images from close distance to the camera [1], [5], which limit the applicability of the systems to be considered for forensic and surveillance applications.

The images acquired using visible imaging under less controlled environments appears to be noisier as compared to the NIR images and therefore require development of new approaches. The visible illumination images are usually influenced by multiple noise sources, such as motion/defocus blur, occlusions from eyelashes, hair and eyeglasses, reflections, off-angle and partial eye images, as shown by some sample images in Figure 1. There have been some promising efforts devoted in developing reliable iris segmentation approaches to segment such noisy images which are acquired using visible imaging [7]-[9], [15]. Further research on this particular area emphasizes on developing robust feature extraction and matching algorithms which can accommodate inherent image variations in the segmented noisy iris images. NICE.II was an open competition with the focus specifically on feature extraction and matching algorithms of the noisy iris images [16] (please see [20] for detailed analysis on the winning algorithms). Apart from that, there exist quite limited literatures in this particular study. Table 1 attempts to summarize recent prominent efforts in the feature extraction and matching techniques for iris recognition. Apparently, none of the existing works have provided rigorous performance evaluation for both the visible and NIR iris images which are distantly acquired from less controlled conditions.

Figure 1: Sample images acquired at-a-distance under less controlled environments using visible illumination imaging.

1.1. Related Work
Most of the existing commercial iris recognition systems employ a matching model which is relies on the binarized (iris) codes, which are generated from the quantization of the filter responses [1]-[2]. Ref. [12] presented the first detailed work in the literature which illustrated the
existence of inconsistent (fragile) bits in the iris codes. Such inconsistency may be attributed to the segmentation error, alignment issues and the choice of the employed filter. Fragile bits are regarded as bits which are not temporally consistent, i.e. the values of some bits are flipping between 0 and 1 across the iris codes which are acquired at different time instance from the same subject. The fragile bits are learned for each subject already enrolled in the system and are therefore considered personalized to each individual registered with the system. In order to determine which bits are fragile, a global threshold must be predetermined. For example, let’s $g_0$ and $g_1$ respectively denote the number of occurrences of bit $n$ equals to zero and one, which are statistically learned from $M$ iris codes ($g_0 + g_1 = M$ ) of the same subject. The bit $n$ is considered to be fragile such that $\frac{|g_0 - g_1|}{g_0 + g_1} \geq \tau$.

Such globally determined threshold can be obtained from the training data, but is not adequate to determine the fragile bits for each distinct subject which may have different levels of fragility in their iris codes. Reference [13] considered a further research effort to overcome such limitation by employing a weight map which is personalized to each of the distinct subjects. The personalized weight maps are derived from the knowledge of the fragile bits and can be computed as follows:

$$P = \frac{1}{M \times M} \sum_{\alpha=1}^{M} \sum_{\beta=1}^{M} \text{IrisCode}_{\alpha} \oplus \text{IrisCode}_{\beta}$$

where $\oplus$ and $\neg$ denote the exclusive-OR and negation Boolean operations respectively. Equation (1) computes the average probability of statistically learned fragile bits in order to obtain a more robust representation of the fragile bits. Due to the negation operation, computed result from (1) is referred as stability map in this paper. The weight map is then obtained by passing the learned stability map through a normalized nonlinear compressive function:

$$w_h = 2 \frac{g_0 - g_1}{g_0 + g_1} - 1.$$  

For performing iris matching, a personalized matching strategy was proposed to quantify the similarity between test iris code and the registered iris code of class $j$, which is given as follows:

$$HD_j = \frac{\| (\text{IrisCode}_{\text{test}} \oplus \text{IrisCode}_j) \times W_j \|}{\| W_j \|}$$

where $W_j$ is the computed weight map specifically for class $j$ or for the iris images from $j^{th}$ subject.

### 1.2. Our Work

Iris recognition from the distantly acquired eye images and under less constrained environments poses several challenges, especially for the images acquired using visible imaging. Image quality degradation is usually higher for the visible illumination iris images acquired from such dynamic environments. Most of the existing iris matching algorithms such as those in [4], [11]-[14], [21] are developed for the NIR iris images and may not perform robustly for the noisy visible illumination iris images. For example, the [12] requires prior determination of a global threshold for masking the fragile bits. As the visible illumination iris images are often acquired under less constrained environments, we can expect high degree of image quality variations across the acquired images, even for those acquired from the same subject. Therefore, it is challenging to predefine such global threshold which can effectively account for such image variations. Approach in [13] employs a personalized weight map strategy to quantify each bit in the iris code according to the consistency of the bits. The computation of the weight map assumes all bits in the iris code are independent and the weights are independently calculated from the stability of each bit. Such assumption may not be appropriate to model the impact of the noise that corrupts iris bits from the iris images acquired at-a-distance, especially under visible illumination. Approach in [14] adopts a sparse representation framework to perform the iris recognition. In sparse representation framework, clear representative iris images are required for the construction of dictionary, such that sparse coefficient can be accurately estimated during the sparse recovery process. As such, we can expect the
influence in the recognition performance when the noisy iris images acquired under less constrained environments are employed. On the other hand, there are very limited efforts in the literature that attempt to address limitations of popular iris matching algorithms on visible illumination images, with notable exceptions like those reported in [5], [20] from NICE.II competition. The winning algorithm from this competition [22] is largely due the usage of multimodal strategy by employing multiple pieces of information from iris, periocular, color distribution, etc. to further improve the recognition performance. In summary, such performance improvement is quite intuitive/expected and can be largely attributed to the multibiometrics strategy, rather than any efforts that solely consider iris information.

Our key objective in this work has been to develop an iris matching strategy which can be more accurate and robust to both NIR and visible illumination iris images which are remotely acquired under less controlled environments. The developed approach is largely motivated from the recent work in [12]-[13], which exploit the consistency of the iris bits from each of the distinct subjects. The accompanying noise distribution in the noisy iris images acquired from such dynamic environments is assumed to be nonlinear or follow a power-law. We further assume that such inherent noise is embedded in the feature space, \textit{i.e.} iris code. In what follows, the phenomenon of the existence of fragile bits can be considered as the iris bits corrupted by noise. Therefore, noise estimation can be regarded as a process of determining the inconsistent bits in the iris codes and there exists a strong correlation between the inherent noise and the stability of the iris bits in the feature space. High degrees of freedom in iris codes have been mathematically proven in [2] to show the uniqueness and randomness of the iris textures. However, the intra-class variations in the iris codes are considered to be minimal, as compared to the inter-class variations. Such small intra-class variations property plays a vital role as exploited in [12]-[13] to determine the fragility of iris bits from the temporal iris codes\footnote{Iris images are acquired in different sessions at a time interval.}. Therefore, the stability map as in (2) reflects the evolution of the iris bits into self-organized critical structures of states which are stable for those bits which are less likely to be degraded by noise. The relationship between the stability and the inherent noise for the iris codes computed from such noisy iris images cannot be simply expressed with a linear model. In this work, such relationship is assumed to be non-linear [11] and modeled as a power-law. The experimental results as reported in section 3 further establish the superiority of the proposed model, which have achieved average improvement of 13.9\% and 47.2\% in the rank-one recognition accuracy as compared to the two most recent competing approaches [12]-[13], on three publicly available databases: CASIA.v4-distance and UBIRIS.v2, respectively. The main contributions of this paper can be summarized as follows:

- This paper has investigated an iris matching strategy which can be effective for both NIR and visible illumination iris images that are distantly acquired under less constrained environments.
- Noise model presented is derived from the stability of temporal iris codes. The relationship between the iris code stability and the accompanying noise in feature space is modeled as based on power-law. The validity or effectiveness of the considered model is ascertained from the rigorous experiments (please see section 3).

The remainder of this paper is organized as follows. In section 2, the developed iris matching approach is detailed. section 3 provides the experimental details and the experimental results. Lastly, discussion and conclusion from this paper are presented in section 4.

2. Proposed Iris Matching Strategy

The block diagram of the iris recognition scheme investigated in this paper are shown in Figure 2. Our work is inspired by promising efforts in [12]-[13] and exploits the fragile bits in the iris codes to further improve the iris recognition performance. The fragile bits are considered as the bits corrupted by noise and are embedded in the feature space. Therefore, there exists a strong relationship between the consistency of iris bits and the embedded noise, which can be effectively modeled using a non-linear relationship or governed by a power-law. The consistency of the iris bits can be learned from the enrolled iris codes for each distinct subject. The mode of the application determines how such learned fragile bits are employed in the matching phase. For the verification (1:1) mode, the learned fragile bits for the claimed identity are employed in generating a single matching score. The computed matching score is then compared with a predefined threshold for decision making. For the recognition mode (1:N), the learned fragile bits from each distinct class are employed and a set of matching scores is generated. An identity is assigned to the test subject from the computed matching score which indicates the most similarity. In this section, we provide further details for the investigated iris matching approach. In particular, the nonlinear noise model can lead to a generalized representation to the iris code. It is to be noted that our present work is focusing on the iris matching phase and prior assumptions are asserted that the iris images are well segmented and normalized. Therefore, all the employed iris images in the experiments are segmented using an improved approach as described in [8], which has been shown to work effectively for at-a-distance face/eye images acquired using both NIR and visible imaging.
2.1. Preprocessing for Normalized Iris Image

The accuracy of iris segmentation and the influence of noise are the primary causes that will definitely influence the recognition performance. The commonly observed noise sources such as the occlusions from the eyelashes, eyelid, hair, eyeglasses and specular reflections in the iris images are usually identified and masked during the iris segmentation process. When the imaging condition is relaxed for remote acquisition of the iris images from less constrained conditions, the acquired iris images are generally appear to be noisier due to the influence from the multiple noise sources. In order to address for such limitation, the normalized iris images are remapped by employing an overlapping blocks strategy as defined as defined in the following:

\[
 f_{b,s} : \mathbb{R}^2 \rightarrow \mathbb{R}^2
\]

(3)

where \( f \) is a function which extracts the image patches of size \( b \times b \) from the normalized iris image, sliding at an interval of \( s \) pixels for both the horizontal and vertical directions. In this work, the interval \( s \) is defined as half of the block size, i.e., \( s = b/2 \). The remapped normalized iris images contain the undesired blocking artifacts caused by the overlapping operation which may introduce the high frequency contents during the iris extraction stage. Therefore a two-dimensional median filter with the size of \( b \times b \) is applied to the remapped normalized iris images to mitigate such blocking artifacts. Such overlapping blocks strategy is empirically observed to achieve better recognition performance on the training dataset for CASIA.v4-distance, as shown in Figure 3. There are two possible reasons to explain why such overlapping blocks strategy is advantageous. Firstly, better localized information redundancy is achieved across each overlapping block. Such redundant information is exploited to better account for the spatial variations in normalized iris images, especially for those acquired under less constrained environments. Secondly, the smoothing operation using the median filter not only mitigates the blocking artifacts, but also simultaneously helps in suppressing the noise in each of the overlapping blocks.

\[
 \text{Recognition Accuracy}
\]

Figure 2: Block diagram of the developed iris recognition scheme. (a) Stability maps learning phase, (b) Matching phase.

2.2. Iris Bits Weighting

The inherent noise in the at-a-distance (noisy) iris images is assumed to be embedded in the iris codes. The phenomenon of the fragile bits [12] provides the evidence such that all such bits available in an iris code are not equally important. By exploiting the temporal intra-class information, the consistency of the bits can be statically learned. Therefore, the bits which are less consistent can be considered as a result of the noise perturbation. Such relationship between the bit consistency and the noise can be modeled using a non-linear function that follows a power-law. Given \( T_j \) training normalized iris images from \( j \)-th class, we first obtain the corresponding iris code representations. Let’s denote \( \theta_n \) as the number of times a \( n \)-th bit is flipping (please see [12] and [13] for further details), then the probability that \( n \)-th bit can flip is computed as follows:

\[
 p_n^j = 1 - \frac{\theta_n}{T_j} \in [0, 1]
\]

(4)

Hence, a stability map \( P^j \) which indicates the consistency of the bits for some iris codes in \( j \)-th class is computed using (1), such that:

\[
 p^j = \{p_1^j, ..., p_N^j\}
\]

A high value of \( p_n^j \) indicates that the respective \( n \)-th bit is expected to be more stable.

Figure 3: Recognition performance by employing overlapping and non-overlapping image blocks.
In order to emphasize those bits which are highly stable, we propose a non-linear mapping function as given as follows:

$$\omega_n^j = \left( p_n^j \right) C^j$$  

(6)

where $C^j$ denotes the adaptive crest factor (peak-to-average ratio) which is employed to measure the quality of the stability map $P_n$ and can be estimated as follows:

$$C^j = \left\{ \begin{array}{ll}
\frac{1}{\mu^j} & \text{if } \mu^j > 0 \\
1 & \text{if } \mu^j = 0
\end{array} \right. $$  

(7)

where $\mu^j = 1/N \sum_{n=1}^{N} p_n^j$ and $p_{peak}$ indicates the maximum probability value. The nonlinear weighting function employed in of work exhibits some desirable properties which can be summarized as in the following:

- The weighting function preserves the local stability values, i.e. when $p_n^j = \{0, 1\}$, regardless of the global stability factor $\mu^j$. As such, the weights for those highly stable (unstable) bits will not be influenced even the $\mu^j$ is low (high).
- When $\mu^j = \{0, 1\}$, which are the two special cases when the global stability factor at its extremum, the weights remain the same, i.e. $\omega_n^j = p_n^j$.
- When $T_n^j = 1$, the computed weight map will be exactly the same as the generated iris code, such that the $\omega_n^j = \{0, 1\}$. Therefore, the equation (6) can be considered as a generalized representation of the conventional iris code representation.

Once the weights are obtained, the similarity between two iris images can be computed from the normalized Hamming distance using (2).

### 3. Experiments and Results

In this section, experimental results obtained from the rigorous experiments on three publicly available databases, namely CASIA.v4-distance [19] and UBIRIS.v2 [18] are reported to further ascertain the superiority of the proposed iris matching algorithm. The three employed databases were distantly acquired (ranging from 3-8 meters) using either visible or NIR imaging from less constrained conditions and such properties are appropriate to meet the objectives of this paper. Since that the focus of this paper is exclusively on the accurate iris matching, all the iris matching algorithms presented in this section employ the identical set of the segmented images which are obtained from the same iris segmentation algorithm.

#### 3.1. Recognition Performance

In order to investigate the performance of the proposed approach, three publicly available databases were employed. In this set of experiments, 1-D log-Gabor filter [17] was employed to extract discriminant iris features from the normalized iris images. The parameters of the log-Gabor filter, i.e. wavelength and SigmaOnf, are exclusively optimized for each employed databases, as can be observed from Table 2. It has to be noted that the images employed in parameters selection are completely independent from those in the test environment. For CASIA.v4-distance database, the images from the first 10 subjects were employed as training images and the first eight left eye images from the rest of 131 subjects were employed as test data. For UBIRIS.v2 database, a subset which consists of 1000 images from 171 subjects as released for [16] was employed. The 96 images from the first 19 subjects were employed for training while the rest was employed as test data (after successful segmentation). For color images, we employed the luma-channel (Y) of the YCbCr after the color space conversion from the RGB color space. As for test environment, we employed the first five images (or at most 5) as the gallery dataset while the remaining images as the test dataset. The images from the gallery were employed in learning the stability maps for each distinct subject.

The cumulative match characteristic (CMC) and the receiver operating characteristic (ROC) curves from the conducted experiments are respectively shown in Figure 4 and 5 to ascertain both the recognition and verification performance of the proposed matching algorithm. These figures also illustrate the corresponding performance from some of the most recent state-of-the-art iris matching algorithms, such as [12]-[13], [14]. The rank-one recognition accuracies are reported as 94.66% and 48.19% on CASIA.v4-distance and UBIRIS.v2 databases, respectively. As compared to the two most recent competing algorithms [13]-[14], the proposed approach achieves average improvement of 13.9% and 47.2% in the rank-one recognition accuracy on CASIA.v4-distance and UBIRIS.v2 databases, respectively. However, the verification performance from the proposed algorithm does not consistently outperform the other competing algorithms, as can be observed from the Figure 5. Such inconsistency can possibly due to the different expectation in the number of gallery images required for the verification.

### Table 2: The employed databases and the log-Gabor parameters.

<table>
<thead>
<tr>
<th>Database</th>
<th>CASIA.v4-distance</th>
<th>UBIRIS.v2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imaging Type</td>
<td>NIR</td>
<td>Visible</td>
</tr>
<tr>
<td>Number of images/Subjects</td>
<td>957/131</td>
<td>864/131</td>
</tr>
<tr>
<td>Size of normalized iris image</td>
<td>512 × 64</td>
<td>512 × 64</td>
</tr>
<tr>
<td>Log-Gabor Parameters (Wavelength / SigmaOnf)</td>
<td>20/0.25</td>
<td>59/0.32</td>
</tr>
</tbody>
</table>

The numbers of iris images available from each of the subjects in UBIRIS.v2 and FRGC databases are not fixed. Some poorly segmented or low-quality images were filtered out by the completely automated segmentation algorithm.

The source code is available at: http://www.umiacs.umd.edu/~jsp/Research/SRRecognition/SparseRecognitionCancelability_PAMI2010.zip

‡ The numbers of iris images available from each of the subjects in UBIRIS.v2 and FRGC databases are not fixed. Some poorly segmented or low-quality images were filtered out by the completely automated segmentation algorithm.
4. Discussion and Conclusions

This paper has investigated a promising iris matching approach for the iris images acquired at-a-distance, using both NIR and visible imaging, under less controlled environments. The accompanying noise in the iris images as acquired from such dynamic environments is more likely to be embedded in the feature space. A noise model is then derived from the temporal information in the feature space by exploiting the consistency and the similarity of the intra-class iris codes. As such, there exists a strong relationship between the stability of the bits in the iris codes and the embedded noise. The stability map as learned from the temporal iris codes provides evidence for those consistent bits which are relatively stable over time. Therefore, the bits which are consistent over time are regarded as distinctive iris features which are less likely to be corrupted by noise, or vice versa. Such relationship between the consistency of iris bits and the embedded noise is modeled using a non-linear function follow a power-law.

The recognition performance achieved by the proposed approach is quite encouraging as the achieved average rank-one recognition accuracies are consistently superior to those achieved from other competing algorithms. The result obtained for the NIR iris images is especially promising but further improvement is required for the more challenging visible illumination iris images. The proposed iris matching approach achieved only comparable verification performance to the [14]. One of the possible reasons may be due to the better feature representation using the sparse representation as the sparse recovery process involves a collective of features from the dictionary. Another possible cause may due to different number of training images required for the verification applications. For that reason, we carried out another experiment on CASIA.v4-distance database to investigate into such matter. We performed the iris segmentation to the rest of the left eye iris images from such database (subject 11-141) using the identical iris segmentation algorithm in order to form a larger image set. After the iris segmentation, the class which has less than 15 images will not be considered in the experiment in order to ensure the same numbers of images are distributed as gallery and test dataset. Therefore, a total of 1200 images from 80 subjects were employed for conducting the experiment. The first 12 images from each subject were reserved as gallery dataset and the remaining 3 was employed for testing. We varied the number of training images using the images from the reserved gallery dataset for computing the stability map. The observation from such experiment is shown in Figure 6(a), which has illustrated different number of training images may be required for the verification. The receiver operating characteristic when 11 images are employed as gallery dataset is also shown in
The experimental results presented in this paper only utilized information from a single biometric modality, i.e., only iris region pixels. A multibiometrics strategy, as also adopted in [21], can provide more promising opportunity to further improve matching accuracy from the acquired eye images. Such an approach to simultaneously use the periocular biometric features has invited increasing attention and some promising results, such as those in [22], have been reported in the literature. The periocular features are often simultaneously acquired with the iris images. Therefore such a strategy that can simultaneously exploit both iris and periocular region information is expected to further improve the matching accuracy and is part of our ongoing further work.

Acknowledgment

This work is partially supported by the competitive research grant from The Hong Kong Polytechnic University (2013-2015), grant no. G-YM09 (PolyU 5175/12E), and PolyU 5011-PPR-12.

References

http://www.csse.uwa.edu.au/~pk/studentprojects/libor