The Best Bits in an Iris Code

Karen P. Hollingsworth, Kevin W. Bowyer, *Fellow*, *IEEE*, and Patrick J. Flynn, *Senior Member*, *IEEE*

Abstract—Iris biometric systems apply filters to iris images to extract information about iris texture. Daugman's approach maps the filter output to a binary iris code. The fractional Hamming distance between two iris codes is computed and decisions about the identity of a person are based on the computed distance. The fractional Hamming distance weights all bits in an iris code equally. However, not all of the bits in an iris code are equally useful. Our research is the first to present experiments documenting that some bits are more consistent than others. Different regions of the iris are compared to evaluate their relative consistency and, contrary to some previous research, we find that the middle bands of the iris are more consistent than the inner bands. The inconsistent-bit phenomenon is evident across genders and different filter types. Possible causes of inconsistencies, such as segmentation, alignment issues, and different filters, are investigated. The inconsistencies are largely due to the coarse quantization of the phase response. Masking iris code bits corresponding to complex filter responses near the axes of the complex plane improves the separation between the match and nonmatch Hamming distance distributions.

Index Terms—Iris biometrics, iris code, texture filter, false reject rate.

1 Introduction

The human iris has extraordinary variations in texture, and it has been shown that iris texture can be used to identify a person. One typical method of using the iris for recognition involves applying a texture filter to an image of the iris and extracting a representation of the texture, called the iris code. The iris code is a set of bits, each one of which indicates whether a given bandpass texture filter applied at a given point on the iris image has a negative or nonnegative result. Many different texture filters and many different lengths of iris code have been investigated, although the Gabor filter proposed by Daugman [2] is considered the traditional approach. Readers interested in other approaches may look at [3] and [4]. In addition, [5] presents a comparative study of different iris recognition algorithms.

For a given iris image, a bit in its corresponding iris code is defined as "fragile" if there is any substantial probability of it ending up a 0 for some images of the iris and a 1 for other images of the same iris. The possibility that fragile bits exist in the iris code was initially suggested by Bolle et al. [6]. They model the theoretical lower bounds for the false accept rate (FAR) and the false reject rate (FRR) of an iris biometrics system and find that "the reported empirical FRR performance degradation is significantly more stable with respect to the system threshold variation than predicted by the theory. This implies that the invariant bits in the iris code representation are dramatically robust to the imaging noise." As a possible explanation for the difference between the empirical and theoretical results, they suggested "... that perhaps not all bits are equally likely to flip, that there are some particularly 'fragile' bits." In this paper, we use the terms "fragile" and "inconsistent" interchangeably, to refer to this phenomenon.

To investigate the existence of fragile bits, we compared iris codes from multiple images of the same eye, for 24 different eyes. This comparison shows that, indeed, some bits are more reliable than other bits in an iris code.

2 Data and Software

The data used for this experiment was collected at the University of Notre Dame and provided to the US National Institute of Standards and Technology (NIST) for use in the Iris Challenge Evaluation (ICE) [7]. All images were acquired using an LG2200 EOU iris imaging system [8]. The subset of the ICE data set used for the experiments in this paper contains over a hundred different images for each of several irises. No more than six images were acquired for a particular iris of a particular subject in any given week. To test the fragility of individual bits in an iris code, we selected images that were mostly unoccluded by eyelids or lashes. Sample images from this data set are shown in Figs. 1, 2, 3, and 4. We selected a subset of the ICE data that contains 24 subjects, with between 15 and 118 images of the left eye of each subject, for a total data set of 1,226 left iris images. This data set contains four Asian and 20 Caucasian subjects. Ten of the subjects are female and 14 are male.

We used software similar to the software IrisBEE [7] to create the iris codes. This software uses 1D log-Gabor wavelets to create a $240 \times 20 \times 2$-bit iris code. The software we used contains improvements to the segmentation as described in [9]. Our software automatically found correct inner and outer boundaries of the iris in all 1,226 images selected for the study (verified by visual inspection). The images were selected to minimize the effects of segmentation errors; however, some minor imperfections in the segmentations could not be avoided. One difficulty in iris

• *The authors are with the Department of Computer Science and Engineering, University of Notre Dame, 384 Fitzpatrick Hall, Notre Dame, IN 46556. E-mail: {kholling, flynn}@nd.edu, kwb@cse.nd.edu.*

Manuscript received 4 Apr. 2008; revised 9 July 2008; accepted 10 July 2008; published online 17 July 2008.
Recommended for acceptance by S. Prabhakar.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-2008-04-0199.
Digital Object Identifier no. 10.1109/TPAMI.2008.185.

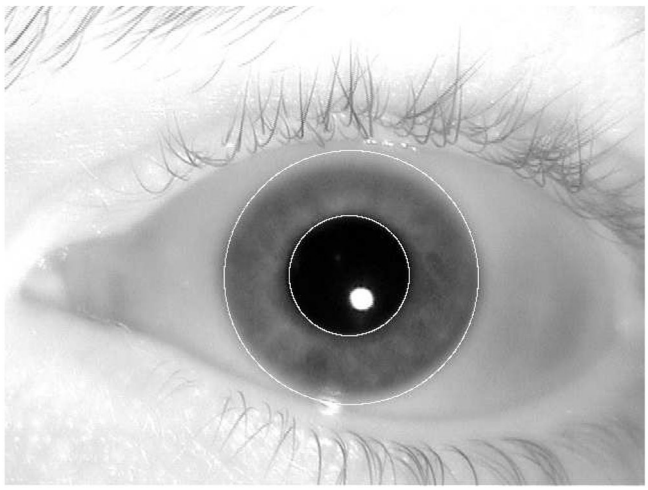

Fig. 1. Sample image 04233d1145 from our data set. This image has no occlusion by eyelids or eyelashes. However, part of the sclera is included in the segmentation of this iris region. Also, some specular highlight is evident in the iris region near the lower eyelid.

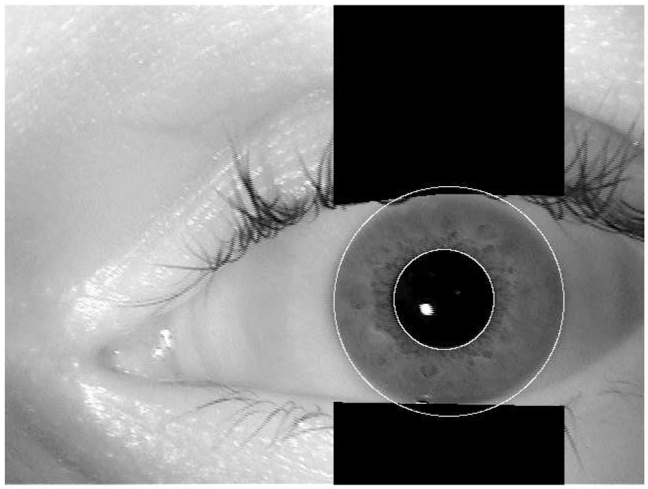

Fig. 3. In this image, parts of the top and bottom of the iris were masked before the iris code was computed. The corresponding parts of the iris code of all images of this iris were dropped from the experimental analysis. This is image 04239d1060 from our data set.

segmentation lies in the fact that iris boundaries are not exactly circular. Therefore, even the best circle caught part of the sclera or eyelid above and below some of the irises (Fig. 1). Also, there are specular highlights on irises of some of the images (Fig. 2). To deal with these segmentation difficulties, we later consider how masking parts of the iris affects our results.

In cases where the segmentation software detected occlusion by eyelids, parts of the iris region were masked (Fig. 3). If part of the iris code was masked in even one image of a subject, the bits from that part of the iris code were left out of all computations in our experiments on that subject. Some of the images have sharp focus, but a few images are less well focused (Fig. 4). Our experiments show that some bits are consistent across all images, even when a few poorly focused images are included in the experiment. This result is consistent with Bolle's assertion that "the invariant bits in the iris code representation are dramatically robust to the imaging noise" [6].

Two different iris codes from the same eye are not necessarily in the same orientation when acquired from the raw images and, therefore, may have to be rotated about the optical axis to be aligned correctly. We wrote a program that would rotate the iris codes so that all codes from the same eye could be aligned and compared for consistency. We used the first iris code as a reference and aligned each subsequent code to the first code. The rotation that gave the minimum fractional Hamming distance was used.

From this set of aligned binary iris code templates, we created a $240 \times 20 \times 2$ matrix, where each entry in the matrix contains the average value for the corresponding bit in the iris code. Half of these entries corresponded to the real part of the output from the texture filtering process, and half of the entries corresponded to the imaginary part. In order to see patterns in the output, we divided each

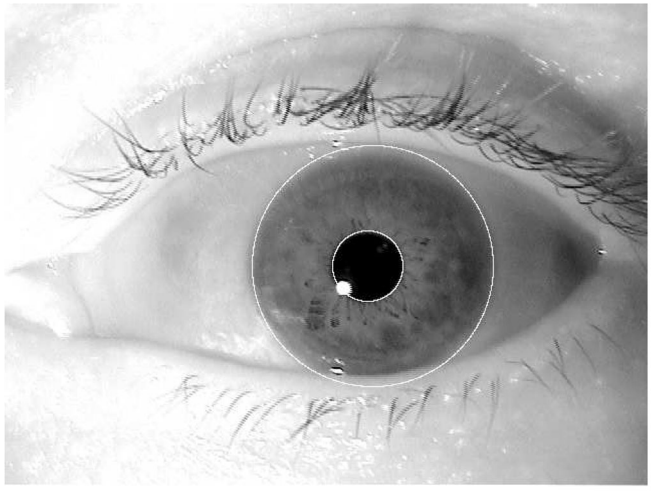

Fig. 2. Sample image 02463d1268 from our data set. This iris contains specular highlights near the bottom of the iris and near the pupil.

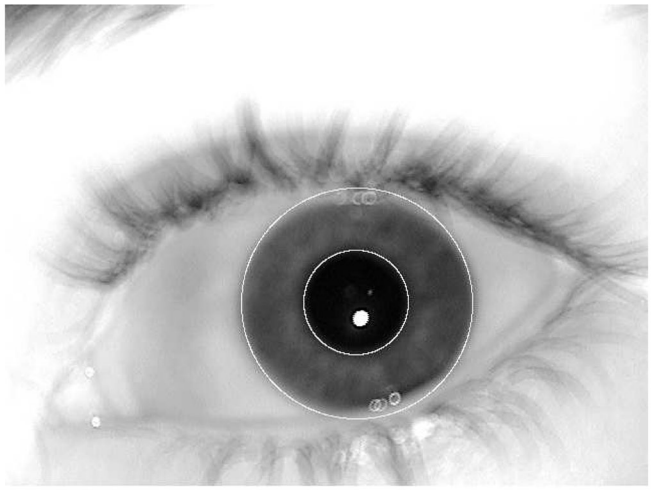

Fig. 4. Despite including some poorly focused images in the data set, on average 15.96 percent of bits in the iris codes were perfectly consistent. This is image 04336d692 from our data set.

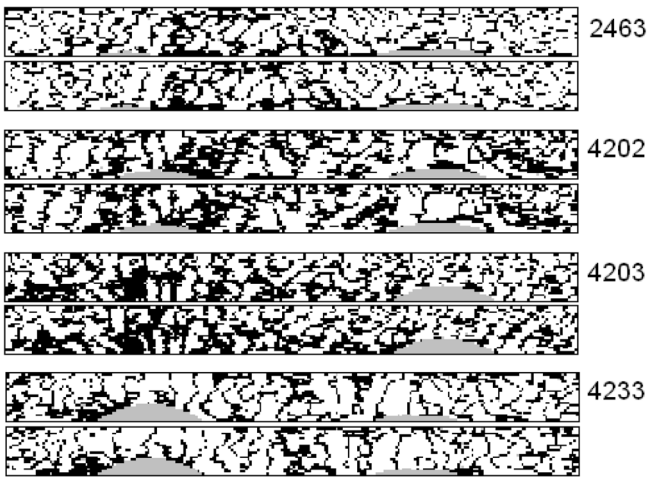

Fig. 5. Black areas in each rectangle are inconsistent parts of the iris code, and white areas are consistent. Each rectangle represents data from the iris codes of at least 15 different images of the same eye. Two rectangles are shown for each subject; one rectangle shows the real bits in the iris code, and the next rectangle shows the imaginary bits. Light gray regions are masked regions. The numbers on the side are the subject numbers associated with images in the ICE data set.

matrix into two $240 \times 20 \times 1$ matrices, where one matrix represented the real bits of the iris code, and the other represented the imaginary bits.

3 Existence of Fragile Bits

All subjects had three different regions apparent in their iris codes: areas consistently equal to 0, areas consistently equal to 1, and inconsistent areas. The inconsistent areas tended to occur at the boundaries between regions of zeros and regions of ones. Examples of the inconsistent regions in the iris codes are shown in Fig. 5. For this particular figure, we arbitrarily defined 30 percent as our threshold for determining inconsistent bits. If a bit was equal to one the majority of the time but was equal to zero 30 percent of time, then we say that the bit "flipped" in 30 percent of the iris codes. If a bit was zero the majority of the time but one for 30 percent of the time, we also say that the bit "flipped" in 30 percent of the iris codes. In this figure, the black regions correspond to bits that were flipped in at least 30 percent of the iris codes. If we are less strict in our definition of what constitutes an inconsistent bit, then there will be a greater number of inconsistent bits. Fig. 6 shows that if we consider

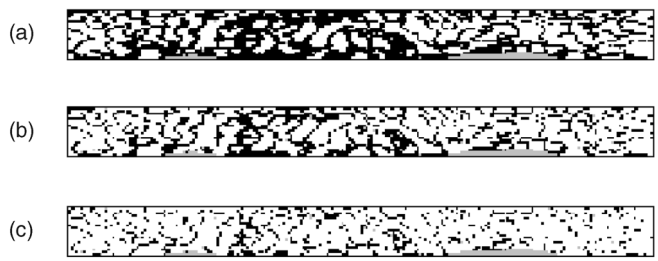

Fig. 6. The number of fragile bits depends on the threshold used. There are fewer bits that flip in at least 40 percent of the iris codes than there are that flip in at least 20 percent of the iris codes. (a) $p = 20\%$. (b) $p = 30\%$. (c) $p = 40\%$.

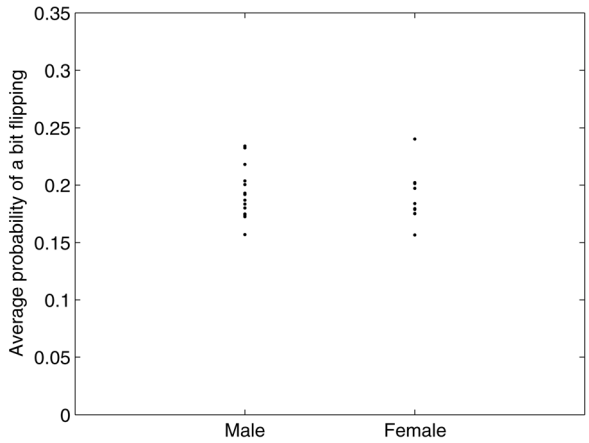

Fig. 7. There is no difference in the average fragility of iris code bits between the two different genders.

any bit that flips more than 20 percent of the time to be fragile, then there will be many fragile bits. If we consider any bit that flips more than 40 percent of the time, then there will be fewer fragile bits. In our study we found that, on average across our set of images, 15.96 percent of the bits in an iris code were perfectly consistent; that is, 15.96 percent of the unmasked bits were always equal to 1 or always equal to 0, for all iris codes for an iris. The sample standard deviation across the 24 irises was 6.39 percent. The subject with the smallest fraction had 4.74 percent of the bits perfectly consistent and the subject with the largest fraction had 33.2 percent of the bits perfectly consistent. One reason why some of our subjects have a greater number of perfectly consistent bits than another is that some subjects have fewer images in our data set than others. As we acquire more images of an iris, there are more opportunities to acquire outliers in any given bit in the iris code.

In addition to considering perfectly consistent bits, we also looked at the average fragility of the unmasked bits of a subject. For each bit, the percentage of the images in which the bit flips must lie between 0 percent and 50 percent. We found the frequency that each bit flipped and took the average across all bits for a subject. This average fragility for each subject is not correlated with the number of images we had for the subject. The correlation coefficient between these factors is only -0.1730, which is not statistically significant (p-value 0.4188).

Fragile bits show up in about equal amounts in both males and females. We use average fragility in making comparisons between genders because average fragility did not appear to be dependent on number of images. Fig. 7 shows the average fragility of the 24 subjects when divided into groups of males and females. We applied a two-tailed Student's t-test to the null hypothesis that the mean average fragility for males was equal to the mean average fragility for females. The alternative hypothesis was that the means were unequal. We found no evidence to reject the null hypothesis (p-value 0.6751), so there appears to be no difference between the two genders.

The spatial pattern apparent in these consistent and inconsistent regions likely comes as a result of how the iris image is processed to generate the code. A Fourier transform

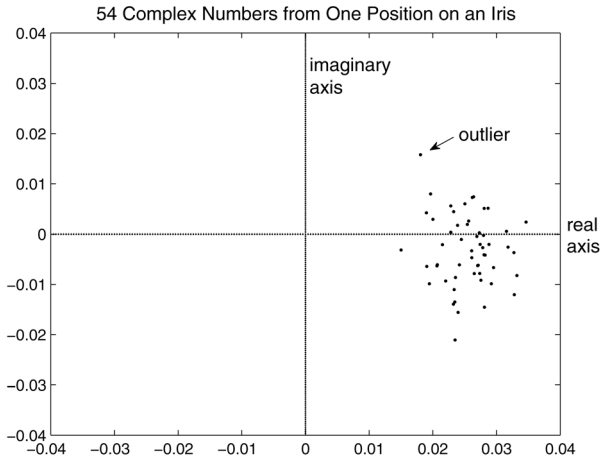

Fig. 8. These 54 complex numbers, each from the same region on 54 different images of the same subject's eye, all correspond to the same location on the iris code. Each complex number is mapped to two bits. This particular part of the iris code had a highly consistent real bit and a highly inconsistent imaginary bit.

is applied to the unwrapped, polar representation of the iris, and then, the values are multiplied by a log-Gabor filter. Next, an inverse Fourier transform is applied, yielding 4,800 complex filter responses in the spatial domain. Rather than storing the complex numbers as the iris codes, the phase of each complex number is quantized to one of the four quadrants. A complex number in the first quadrant of the complex plane is mapped to the value, 11; the second quadrant, 01; the third quadrant, 00; and the fourth quadrant, 10. If a region of the iris image were associated with a complex number near the negative imaginary axis, small changes in the complex number could make that region of the iris map to a 00 some of the time, and a 10 at other times. In this case, we would expect the real bit to be fragile because half of the time that bit would be a 0, and half of the time that bit would be a 1. Furthermore, we would expect the imaginary bit to be consistent, because the imaginary bit is equal to zero no matter whether the complex number is in the third or fourth quadrant.

Fig. 8 shows an example of the distribution of 54 complex numbers from 54 different images of the same iris. Each of these 54 complex numbers is associated with the same location in the iris code. In particular, this location on the iris code had a highly inconsistent imaginary bit and a highly consistent real bit. As we expected, the complex numbers associated with these two bits of the iris lay close to the positive real axis.

4 Outliers

We examined the distributions of complex numbers for multiple subjects and multiple positions on the iris. We found that many of these distributions had outliers (for example, see Fig. 8). For one particular subject (subject 02463), we manually marked and examined all outliers that fell more than four sample standard deviations away from the mean in their data set.

TABLE 1
Sources of Outliers

Source	Percent of Outliers due to this Source
Light regions	85.9% (1270 of 1478)
Dark borders around specular highlights	11.8% (174 of 1478)
Shadows	2.3% (34 of 1478)

The majority of the outliers were due to spots on the iris that were lighter than the surrounding region (Table 1). These light spots seem to be specular highlights. Some of the brightest specular highlights have a dark border surrounding them. The dark borders also cause outliers in the data. Fig. 9 displays a portion of a sample iris image showing outliers caused by specular highlights and dark borders.

There are surfaces in the eye that reflect light and thus create the highlights that are causing outliers. For instance, light can be reflected off of the 1) outer and 2) inner surfaces of the cornea and off the 3) outer and 4) inner surfaces of the lens of the eye. These four types of reflections are called Purkinje images [10]. Since the lens is located behind the iris, the light regions seen on the iris are probably not reflections from the lens, but they could be reflections from the cornea. In some instances, there appears to be faint specular reflections close to and in line with brighter specular reflections. Fig. 10 shows examples of these types of highlights.

While most outliers occurred on or near light regions of the iris image, some outliers appeared on shadows in the image that were not near any highlights. We surmise that these outliers are due to shadows on the iris or perhaps

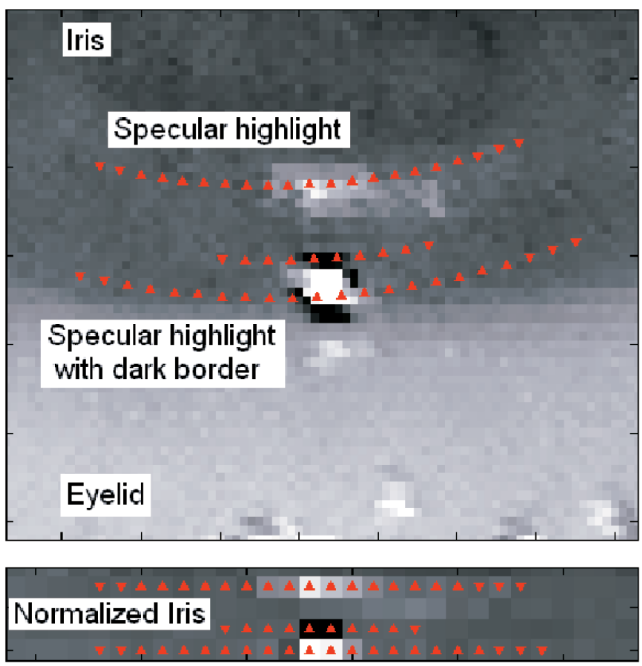

Fig. 9. Close-up view of image 02463d1323 showing specular highlights. Triangles mark the positions on the iris that resulted in outliers after the 1D log-Gabor filter was applied. A zoomed-in portion of the corresponding normalized iris is also shown.

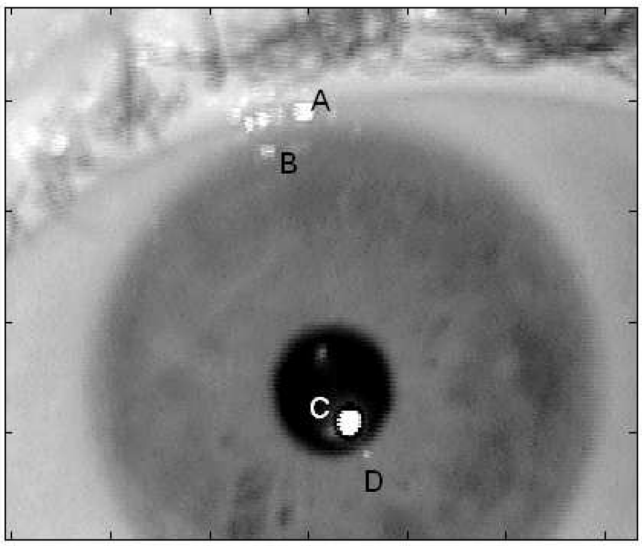

Fig. 10. Faint reflections often appear close to brighter reflections. This image contains (A) bright reflections on the eyelid, (B) a nearby, faint reflection on the iris, (C) a bright reflection in the pupil, and (D) a nearby, faint reflection on the iris. Both of the faint reflections marked in this image (B and D) caused outliers in the complex data (image 02463d1329).

features that are small enough that they only show up in the best-focused images. Fig. 11 shows an example of this type of outlier.

5 Inner versus Outer Radial Bands of the Iris

Several researchers [11], [12], [13] have suggested using only the inner part of the iris for recognition. Du et al. [14] suggested that "a more distinguishable and individually

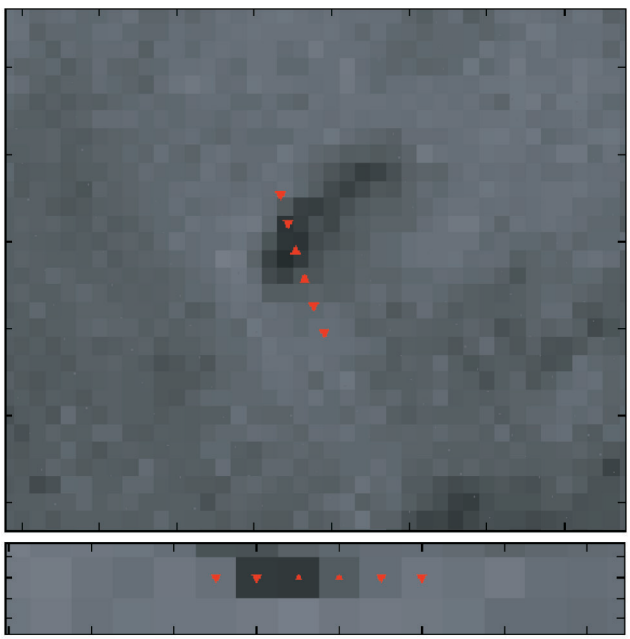

Fig. 11. Some outliers showed up on dark regions of the iris image (02463d1276).

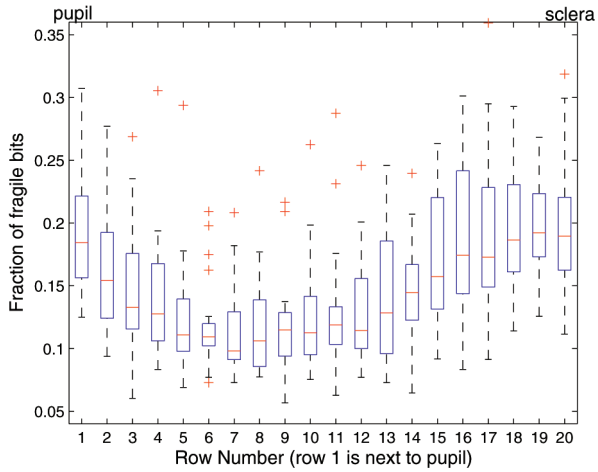

Fig. 12. This figure shows the percent of fragile bits in each row of the iris code. Rows in the middle of the iris code (rows 5 through 12) are the most consistent.

unique signal is found in the inner rings of the iris" and that "as one traverses to the limbic boundary of the iris, the pattern becomes less defined and ultimately less useful in determining identity." Since the inner part of the iris is less likely to be occluded, we suspect that it is the occlusion, rather than the lack of texture, that might make outer bands less valuable. To test this idea, we graphed the percent of fragile bits that occurred at each radial band in the iris.

Using a consistency threshold of 40 percent, we calculated the percent of bits that were fragile for each row of the "unwrapped" iris image (corresponding to a radial band), for all 24 iris images. We found that the middle rows of the iris were the most consistent (Fig. 12). Our experiment yields the same conclusion as that of another contemporary experiment [15].

Despite our efforts to obtain unoccluded images for the data set, we suspected that specular highlights close to the eyelids could be affecting our measurements of the consistency of the different bands of the iris code. Therefore, we applied a mask to the data to disregard the upper and lower quarters of the iris and then we looked at the percent of fragile bits for each row of the iris code in the unmasked quarters of the iris (left and right sides of the iris). The mask used is shown in Fig. 13, and the result is shown in Fig. 14.

We also performed the same test using fragility thresholds of 30 percent and 20 percent. In all cases, masking the top and bottom quarters improved the consistency of the outer rows. Clearly, the quality of the segmentation and highlight detection affects the value of the outer rows of the iris. Another possible reason for some of the inconsistencies is that our software uses circles in modeling the iris boundaries. If we allowed noncircular boundaries, the inner and outer rows might be more consistent. Dilation is another factor that could affect consistency, especially in the rows closest to the pupil [15].

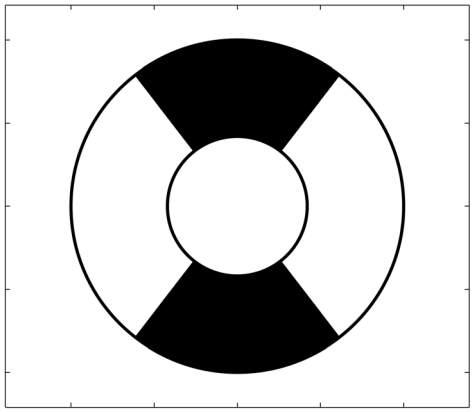

Fig. 13. Mask used for evaluating the consistency of the inner versus outer bands of the iris. When we consider only the right and left sides of the iris and disregard the top and bottom sections, which had the most specular highlights and occlusion, there is not as much spread in the fraction of fragile bits across the rows. There still seems to be a high percentage of fragile bits in rows 1 and 2 of the iris code; however, the general trend shows that all rows of the iris code have a high amount of consistent information.

6 Effect of Granularity of Iris Alignment on Consistency

Since we have experimentally shown the existence of fragile bits, it is reasonable to ask how small modifications in the iris recognition algorithm might affect the consistency of the bits in the iris code. One important requirement of an iris recognition algorithm is that the algorithm must be rotation invariant. That is, a small tilt of the head should not cause the recognition algorithm to fail. Daugman [2] suggested that to achieve rotation invariance, an enrolled iris code could be compared to several different shifts of the probe iris code and the shift that yielded the smallest fractional Hamming distance could be taken as the correct orientation of the probe image.

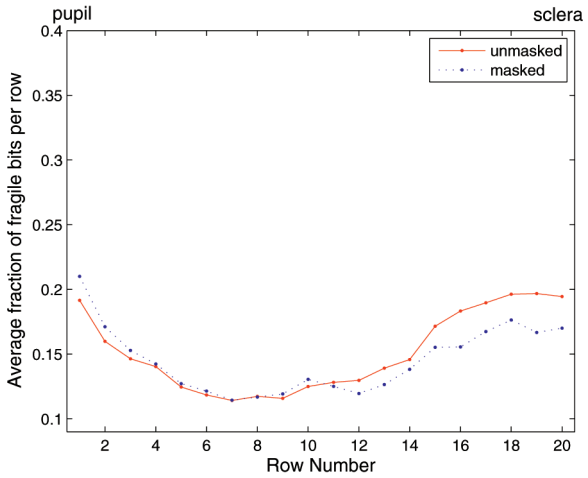

Fig. 14. When the regions most affected by specular highlights and occlusion are masked, the consistency of the outer rows of the iris code improves.

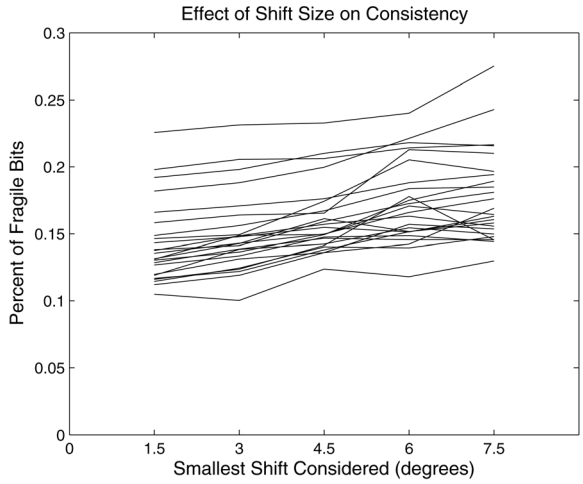

Fig. 15. In matching a pair of iris codes, multiple different orientations of the probe iris code are considered. Allowing for smaller possible rotations decreases the percent of fragile bits.

Our software shifts the iris code by two angular steps at a time by default, which is equivalent to rotating the original iris image by 3 degrees. In order to be as thorough as possible, we used shifts of the iris code as small as one step (1.5 degrees) when testing for the existence of fragile bits. We subsequently needed to check whether changing the shift size would affect the consistency of the bits.

We tested for the existence of fragile bits using shift sizes as small as 1.5 degrees and as large as 7.5 degrees. The general trend showed that using a finer resolution tended to yield a larger number of consistent bits. However, if we extrapolate on the data, it is clear that the y-intercept of the graph is greater than zero, implying that fragile bits exist regardless of shift resolution. Therefore, even if we could have an infinitesimal shift size, there would still be fragile bits. A graph showing the effect of shift size is shown in Fig. 15. Each line in the graph represents one of our 24 subjects. As in the previous section, we used a threshold of 40 percent to determine which bits were fragile.

7 Effect of Filter on Consistency

Our software uses 1D log-Gabor wavelets in creating the iris code, but there are many different types of filters that can be used in an iris recognition algorithm [16]. We explored the effect of a different filter on the fragile bit patterns in the iris code.

In order to try a different filter implementation, we obtained an open-source iris recognition system, OSIRIS [17]. For segmentation, we used the centers and radii of the irises and pupils generated by our own software. OSIRIS does not currently mask eyelids and eyelashes. However, the images in this data set were selected because they did not contain much occlusion due to eyelids. In addition, in creating the iris code, we chose to ignore the 20 percent of the iris closest to the limbic boundary, instead using sample points closer to the pupil so we would avoid the regions affected by the eyelid (this restriction applies to this section of the paper only).

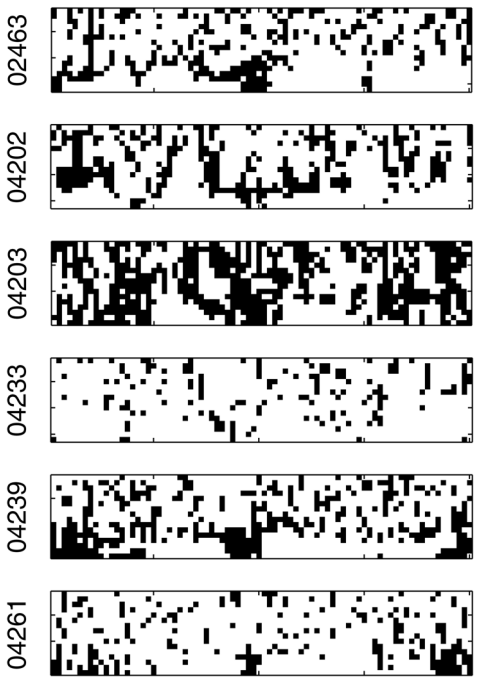

Fig. 16. Each rectangle shows the fragile bits from part of the OSIRIS iris code. Black bits represent fragile bits. The rectangles shown here are the part of the iris code created with the imaginary part of the largest Gabor filter used.

To create the iris code, OSIRIS first processes the segmented iris image to yield a normalized image that is 64 by 512. Next, OSIRIS applies 2D Gabor filters to selected sample points in this image. The filter bank that comes with OSIRIS contains six filters, three for real parts and three for imaginary parts. The filters come in three different sizes: 9 by 15, 9 by 27, and 9 by 51. We selected 1,280 sample points in a 16 by 80 grid pattern on the 80 percent of the normalized image closest to the pupil. Therefore, the resultant iris code is $80 \times 16 \times 6$. Once again, we graphed which parts of the iris code were fragile. As in Fig. 5, the black regions correspond to bits that were flipped in at least 30 percent of the iris codes. Since OSIRIS uses three filters, and each filter has real and imaginary parts, we have six rectangles to display the fragile bits for each iris. The fragile bit patterns for the imaginary parts of the largest filter are shown in Fig. 16 for a few subjects. Clearly, the phenomenon of fragile bits is apparent even when a different filter is used.

In Section 3, we remarked that inconsistent areas in our iris codes tended to occur at the boundaries between regions of zeros and regions of ones. The same pattern is apparent in many of the iris codes for the large filter of OSIRIS. One example of this phenomenon is illustrated in Fig. 17. In this figure, areas consistently equal to 1 are marked in red and areas consistently equal to zero are marked in blue. The yellow and green areas represent inconsistent areas, that are 1s in some of the iris codes for the subject and 0s in other of the iris codes for the subject. The yellow regions clearly are sandwiched between an area of red on one side and blue on the other. Interestingly, this trend is not apparent when a smaller filter is used. We suspect that this trend would appear if we selected sample points that were closer together.

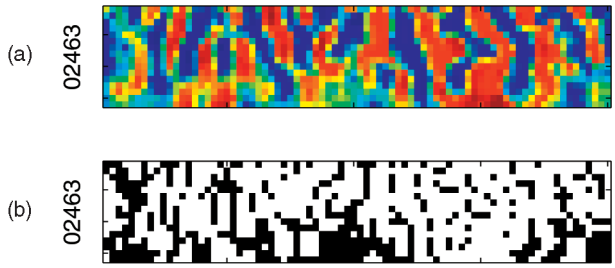

Fig. 17. For the large filter, inconsistent regions of the iris code (shown in yellow) often fall in between regions consistently equal to one (shown in red) and regions consistently equal to zero (shown in blue). In the lower pane, the corresponding black and white figure is shown, with inconsistent regions drawn in black.

Our initial impressions of the fragile bit maps were that the smaller filters had fewer consistent bits than the larger filters. To test this idea, we made histograms showing the consistency of the bits for each of the six filters, across all images of all subjects. These histograms are displayed in Fig. 18. The x-axis shows the percent of times that bits flipped. Bits that fell in the first bin are the most consistent and flipped between 0 and 5 percent of the time. Bits that fell in the last bin are the most fragile and flipped between 45 and 50 percent of the time. The largest filters had the highest number of bits in the first bin and, therefore, in general, the largest filters seem to produce the most consistent bits.

8 Theoretical Impact of Fragile Bits on False Reject Rate

Knowing which bits in a particular subject's iris code are fragile could improve recognition performance. At

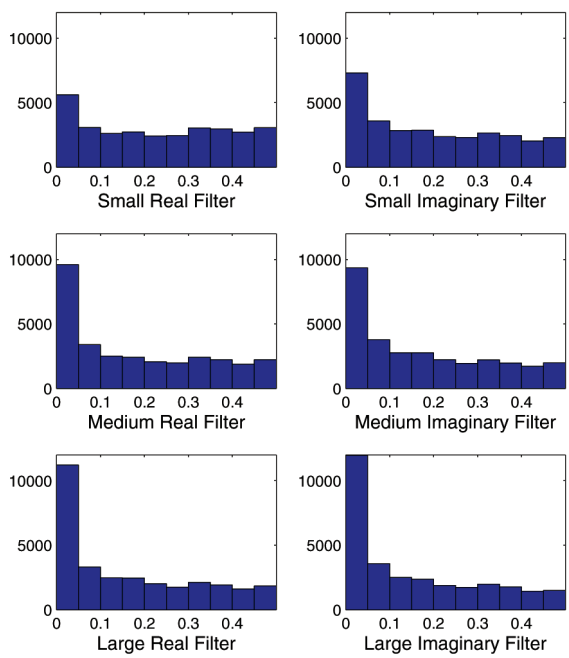

Fig. 18. The largest filters had the highest number of bits in the first bin and, therefore, in general, the largest filters seem to produce the most consistent bits.

enrollment, a sequence of iris images could be taken and analyzed to create the enrollment iris code and a mask that would mask out inconsistent bits. Such a mask could be used in addition to a mask for masking out eyelids and eyelashes. This section presents the theoretical FRR for an iris recognition scenario and then shows how this rate can be improved by masking out the fragile bits in the iris code. We try to follow as closely as possible the notation presented by Bolle et al. [6] in their calculations of FRR. We make two different calculations for FRR. The first calculation presents the error rate for a traditional method that uses all bits in the iris code, and the second calculation presents the error rate for a method that masks fragile bits and only uses the more consistent bits. Let Q and R be "ground truth" iris codes of length N, both from the same iris. When iris code Q is calculated by the iris acquisition system, some of the bits flip so that the result, $\hat{Q}$, is only an approximation to the true iris code, and does not match iris code R exactly. If enough of the bits flip so that the Hamming distance $h(\hat{Q}, R)$ exceeds the decision threshold, d_T, then a false reject occurs. In this instance, the "Hamming distance" is the number of bits that disagree, rather than the fraction of bits that disagree. Bolle et al. [6] showed that the FRR is

$$\begin{aligned} FRR(d_T) &= Pr\Big(h(\hat{Q}, R) > d_T | h(Q, R) = 0\Big) \\ &= \sum_{i=d_T+1}^{N} \binom{N}{i} p^i (1-p)^{N-i}, \end{aligned} \tag{1}$$

where p is the probability that a bit will flip and $(1-p)$ is the probability of not flipping a bit. Since we showed that not all bits have the same probability of flipping, we modify Bolle et al.'s equation to include two different probabilities. Suppose that k of the bits have probability p_1 of flipping, and $N-k$ have probability p_2 of flipping. The first k bits will be termed "Set 1" and the second k bits are "Set 2." Let $p_1 > p_2$ so that Set 1 contains the more fragile bits. Let i be the number of bits that flip when iris code $\hat{Q}$ is acquired. The fractional Hamming distance is the fraction of bits in the two iris codes that disagree, so the probability of a false reject is the probability that $i/N > d_T$ or, equivalently, that $i > N \cdot d_T$. Let j be the number of bits from Set 1 that flip and let $i-j$ be the number of bits from Set 2 that flip. The probability of j bits flipping, out of k bits from Set 1, is

$$\binom{k}{j} (p_1)^j (1-p_1)^{k-j}. \tag{2}$$

The probability of $i-j$ bits flipping, out of the $N-k$ bits in Set 2, is

$$\binom{N-k}{i-j} (p_2)^{i-j} (1-p_2)^{N-k-i+j}. \tag{3}$$

Thus, the total FRR when using all N bits from the iris code is

$$\sum_i \sum_j \binom{k}{j} \binom{N-k}{i-j} (p_1)^j (1-p_1)^{k-j} (p_2)^{i-j} (1-p_2)^{N-k-i+j}, \tag{4}$$

summed over all possible values for i and j: i ranges from $N \cdot d_T$ to N and j ranges from $max\{i-N+k, 0\}$ to $min\{i, k\}$.

Now, consider the situation where the k fragile bits from Set 1 are masked out. Let i^* be the number of remaining bits that flip. The probability of a false reject is the probability that $i^*/(N-k) > d_T$. The total FRR when using $N-k$ bits is

$$\sum_{i^*=(N-k)d_T}^{N-k} \binom{N-k}{i^*} (p_2)^{i^*} (1-p_2)^{N-k-i^*}. \tag{5}$$

To compare the situation of using all N bits to the second situation of using only $N-k$ bits, we assume some numerical values. Bolle et al. [6] used $N = 173$ because Daugman's 1993 paper [18] found 173 independent degrees of freedom in the iris code. In order for our results to be easily comparable to Bolle et al., we also take $N = 173$. We analyzed the experimental consistency data so we could choose reasonable values for k, p_1, and p_2. For each subject, we divided the bits into two groups: fragile bits with probability greater than 0.4 of flipping and consistent bits with probability less than 0.4 of flipping. We counted how many bits fell into each group. On average, 15.0 percent of bits had probability greater than 0.4 of flipping and 85.0 percent of bits had probability less than 0.4 of flipping. The probability of a bit from the first group flipping averaged 0.4483 and the probability of a bit from the second group flipping averaged 0.1467. Fifteen percent of 173 is approximately 26, so we let $k = 26$. We also let $p_1 = 0.4483$ and let $p_2 = 0.1467$. Assuming these numerical values, the FRR for the traditional scenario is 1.31×10^{-5} and the FRR for the scenario that masks the fragile bits is 3.33×10^{-8}. Thus, detecting fragile bits at enrollment and masking them out at matching could reduce the FRR by three orders of magnitude.

9 Empirical Evidence of Improved Accuracy

To test the situation described in the previous section, we would have to select some subset of images from each subject and assign them to be "enrollment images" while using the remaining images as "probe images." Alternatively, we could try to predict which bits in the iris code would be fragile, by using our knowledge that complex numbers near the axes of the complex plane yielded inconsistent bits in the iris code. Such an approach would only be an approximation to truly detecting and masking fragile bits, but it would be simpler because it would only require one iris image to decide which bits to mask.

In Section 3, we demonstrated that complex numbers near the imaginary axis of the complex plane resulted in fragile real bits, and complex numbers near the real axis resulted in fragile imaginary bits. This idea suggests that one simple way to mask out fragile bits is to mask real bits from complex numbers too close to the imaginary axis and mask imaginary bits too close to the real axis.[1] We ran two recognition experiments; the first experiment used our normal iris recognition software and, in the second experiment, we masked bits close to the axes.

1. Although we have not found documentation in the published literature, we understand that publicly deployed iris recognition algorithms already use this idea.

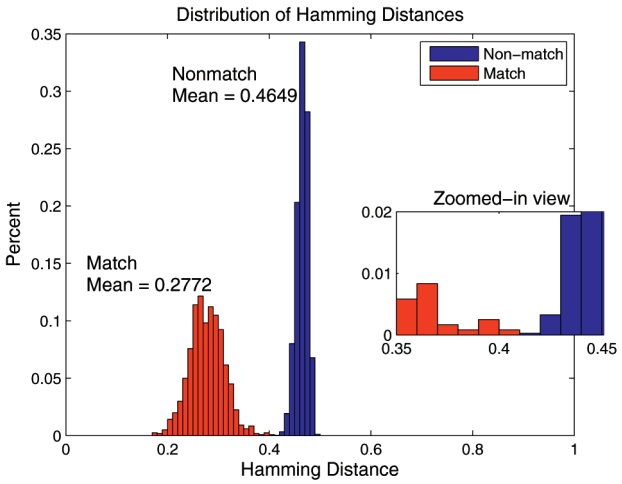

Fig. 19. This figure shows the match and nonmatch distributions for our software, without any masking of fragile bits.

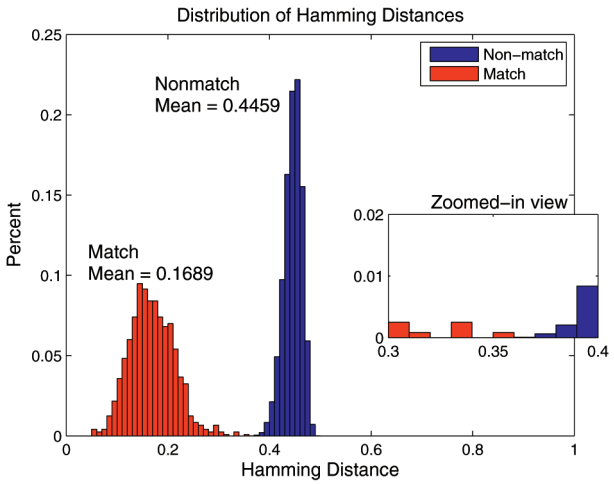

Fig. 20. This figure shows the match and nonmatch distributions for our software when bits close to the complex axes are masked. The match distribution has moved a significant amount to the left, closer to 0, as desired. In addition, the nonmatch distribution has also widened slightly.

For the gallery in our experiment, we took the first image from each subject. The remaining images for each subject were used as the probes. Since we have 24 subjects, we had a total of 24 images in our gallery and 1,202 images in our probe. For both experiments, we had 100 percent rank-one recognition. Therefore, in order to distinguish between the two experiments, we graphed the histograms of the match and nonmatch distributions for each experiment. The experiment resulting in a better separation between the match and the nonmatch distributions is the better method. The result of the first experiment, with no masking of fragile bits, gives the histogram shown in Fig. 19. The mean fractional Hamming distance for all match comparisons was 0.2772 and the mean fractional Hamming distance for all nonmatch comparisons was 0.4649.

For the second experiment, we modified the code for creating the mask of our iris code templates. We took the real parts of all 4,800 complex numbers for the image, took their absolute value, and then sorted them. Next, we identified all the numbers in the lowest quartile of this set. For each complex number with its real part in the lower quartile of the data, we masked the corresponding real bit in the iris code. Finally, we applied the same procedure to sort and mask the lower quartile of the imaginary numbers. This procedure had the effect of masking all real bits close to the imaginary axis of the complex plane, and masking all imaginary bits close to the real axis of the complex plane.[2] The result of this second experiment yields the histogram shown in Fig. 20. The mean fractional Hamming distance for all match comparisons was 0.1689 and the mean fractional Hamming distance for all nonmatch comparisons was 0.4459.

The nonmatch distribution grew slightly wider in the second experiment. The mean of the nonmatch distribution has shifted slightly to the left, by 0.0190. However, the match distribution has also shifted a significant distance to the left. The mean of the match distribution has decreased by 0.1083. In the first experiment, the distance between the two means was 0.1877 and the distance between the two means in the second experiment was 0.2770, a large improvement over the previous performance.

10 Conclusions

The consistency of the different bits in an iris code has not been studied in any other previously published work. Our experiments prove the existence of fragile bits and show that the fragile-bit phenomenon is evident across genders and different filter types. There appears to be no gender difference in the consistency of iris code bits. There does seem to be a difference in consistency of iris code bits based on the size of filter used.

Contrary to some conventional wisdom in the iris biometrics field, we find no significant difference in the value of the inner rings of the iris versus the outer rings. Perhaps surprisingly, our results indicate that the middle bands may be slightly better than either the inner or the outer bands. As noted earlier, this result could arise from the fact that our segmentation process uses enforced circular boundary models.

We show that there are "outliers" in the distribution of values underlying a given iris code bit and that these are largely due to specular reflections. Some of these specular reflections are faint and have generally been overlooked; however, they have a noticeable, negative impact on the consistency of those regions of the iris code.

We present a theoretical argument showing that masking out fragile bits can reduce the FRR in an iris code system by several orders of magnitude. Subsequently, we present an experiment that masks many of the fragile bits in the iris code. This modification of the algorithm significantly increases the separation between the match and nonmatch distributions.

In the future, iris biometrics could potentially be used with extremely large populations. Any application of iris biometrics on a nationwide scale would necessitate extremely

2. We would like to thank J. Daugman for the idea of masking the lower quartile of numbers (as opposed to trying to tune our program using a specific cutoff threshold). From personal communication with Daugman, we understand that he has previously developed and used techniques described in this section.

low error rates. Based on our numerical evaluations, we expect that we could reduce the FRR by three orders of magnitude by using our knowledge of consistent and inconsistent bits.

ACKNOWLEDGMENTS

The iris biometrics research at the University of Notre Dame is supported by the US National Science Foundation under Grant CNS01-30839 and by the US Central Intelligence Agency. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of our sponsors. This is an extended and revised version of a paper presented at the 2007 IEEE Conference on Biometrics: Theory, Applications, and Systems.

REFERENCES

[1] K. Hollingsworth, K. Bowyer, and P. Flynn, "All Iris Code Bits Are Not Created Equal," *Biometrics: Theory, Applications, and Systems,* Sept. 2007.
[2] J. Daugman, "How Iris Recognition Works," *IEEE Trans. Circuits and Systems for Video Technology,* vol. 14, no. 1, pp. 21-30, 2004.
[3] R.P. Wildes, "Iris Recognition," *Biometric Systems: Technology, Design and Performance Evaluation,* pp. 63-95, Spring-Verlag, 2005.
[4] K.W. Bowyer, K.P. Hollingsworth, and P.J. Flynn, "Image Understanding for Iris Biometrics: A Survey," *Computer Vision and Image Understanding,* 2007.
[5] P.J. Phillips, W.T. Scruggs, A.J. O'Toole, P.J. Flynn, K.W. Bowyer, C.L. Schott, and M. Sharpe, "FRVT 2006 and ICE 2006 Large-Scale Results," Technical Report NISTIR 7408,, Nat'l Inst. Standards and Technology, http://iris.nist.gov/ice, Mar. 2007.
[6] R.M. Bolle, S. Pankanti, J.H. Connell, and N. Ratha, "Iris Individuality: A Partial Iris Model," *Proc. 17th Int'l Conf. Pattern Recognition,* pp. II: 927-930, 2004.
[7] *Iris Challenge Evaluation 2005 Workshop Presentations,* Nat'l Inst. Standards and Technology, http://iris.nist.gov/ice/presentations.htm, 2008.
[8] LG, http://www.lgiris.com/, Mar. 2009.
[9] X. Liu, K.W. Bowyer, and P.J. Flynn, "Experiments with an Improved Iris Segmentation Algorithm," *Proc. Fourth IEEE Workshop Automatic Identification Technologies,* pp. 118-123, Oct. 2005.
[10] A.J. Glenstrup, "Eye Controlled Media: Present and Future State," master's thesis, Univ. of Copenhagen, http://www.diku.dk/~panic/eyegaze/article.pdf, 1995.
[11] C.-L. Tisse, L. Martin, L. Torres, and M. Robert, "Person Identification Technique Using Human Iris Recognition," *Vision Interface,* pp. 294-299, 2002.
[12] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Local Intensity Variation Analysis for Iris Recognition," *Pattern Recognition,* vol. 37, no. 6, pp. 1287-1298, Feb. 2004.
[13] K. Miyazawa, K. Ito, T. Aoki, K. Kobayashi, and H. Nakajima, "An Efficient Iris Recognition Algorithm Using Phase-Based Image Matching," *Proc. Int'l Conf. Image Processing,* pp. II:49-52, 2005.
[14] Y. Du, B. Bonney, R. Ives, D. Etter, and R. Schultz, "Analysis of Partial Iris Recognition Using a 1-D Approach," *Proc. Int'l Conf. Acoustics, Speech, and Signal Processing,* vol. 2, pp. ii:961-964, 2005.
[15] R. Broussard, L. Kennell, and R. Ives, "Identifying Discriminatory Information Content within the Iris," *Proc. SPIE,* vol. 6944, pp. 69440:T1-T11, 2008.
[16] J. Thornton, M. Savvides, and B.V.K. Vijaya Kumar, "An Evaluation of Iris Pattern Representations," *Biometrics: Theory, Applications, and Systems,* Sept. 2007.
[17] E. Krichen, A. Mellakh, S. Salicetti, and B. Dorizzi, *OSIRIS (Open Source for IRIS) Reference System,* BioSecure Project, http://www.cilab.upf.edu/biosecure1/public_docs_deli/ BioSecure_Deliverable_D02-2-2_b4.pdf.pdf, 2008.
[18] J. Daugman, "High Confidence Visual Recognition of Persons by a Test of Statistical Independence," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 15, no. 11, pp. 1148-1161, Nov. 1993.

Karen P. Hollingsworth received the BS degrees (valedictorian) in computational math and math education from Utah State University in 2004 and the MS degree in computer science and engineering from the University of Notre Dame in 2008. She is a graduate student studying iris biometrics at the University of Notre Dame. She has taught algebra and trigonometry at both the high school and college levels.

Kevin W. Bowyer received the PhD degree in computer science from Duke University. He currently serves as a Schubmehl-Prein Professor and chair of the Department of Computer Science and Engineering, University of Notre Dame. His recent research activities focus on problems in biometrics and in data mining. Particular contributions in biometrics include algorithms for improved accuracy in iris biometrics, face recognition using 3D shape, 2D and 3D ear biometrics, advances in multimodal biometrics, and support of the government's Face Recognition Grand Challenge, Iris Challenge Evaluation, Face Recognition Vendor Test 2006, and Multiple Biometric Grand Challenge programs. His paper "Face Recognition Technology: Security Versus Privacy," published in *IEEE Technology and Society*, was recognized with an "Award of Excellence" from the Society for Technical Communication in 2005. His data mining research has been supported by Sandia National Laboratories. This work focuses on classifier ensemble techniques for problems that exhibit "extreme" characteristics, such as a high imbalance between classes, unusually large size of training data, and noise in the class labels of the training data. Following a year at the Institute for Informatics, ETH, Zurich, he joined the Department of Computer Science and Engineering, University of South Florida (USF). While at USF, he won three teaching awards, received a Distinguished Faculty Award for his mentoring work with underrepresented students in the McNair Scholars Program, and was awarded a sequence of five US National Science Foundation (NSF) site grants for Research Experiences for Undergraduates. In addition, he created the textbook *Ethics and Computing* and led a series of NSF-sponsored workshops on curriculum development in this area. Also during this time, he served as the editor-in-chief of the *IEEE Transactions on Pattern Analysis and Machine Intelligence*, recognized as the premier journal in its areas of coverage, and was elected a fellow of the IEEE for his research in object recognition. He is the founding general chair of the IEEE International Conference on Biometrics Theory, Applications, and Systems (BTAS).

Patrick J. Flynn received the PhD degree in computer science from Michigan State University in 1990. He is a professor of computer science and engineering and a concurrent professor of electrical engineering at the University of Notre Dame. He has held faculty positions at Washington State University and Ohio State University. His research interests include computer vision, biometrics, and image processing. He was an associate editor and an associate editor-in-chief of the *IEEE Transactions on Pattern Analysis and Machine Intelligence*. He is a senior member of the IEEE and a fellow of the International Association for Pattern Recognition (IAPR).