Using Fragile Bit Coincidence to Improve Iris Recognition

Karen P. Hollingsworth, Kevin W. Bowyer, and Patrick J. Flynn

Abstract— **The most common iris biometric algorithm represents the texture of an iris using a binary iris code. Not all bits in an iris code are of equal value. A bit is deemed *fragile* if it varies in value across iris codes created from different images of the same iris. Previous research has shown that iris recognition performance can be improved by masking these fragile bits. Rather than ignoring fragile bits completely, we consider what beneficial information can be obtained from the fragile bits. We find that the locations of fragile bits tend to be consistent across different iris codes of the same eye. We present a metric, called the *fragile bit distance*, which quantitatively measures the coincidence of the fragile bit patterns in two iris codes. We find that score-fusion of fragile bit distance and Hamming distance works better for recognition than Hamming distance alone. This is the first and only work that we are aware of to use the coincidence of fragile bit locations to improve the accuracy of matches.**

I. INTRODUCTION

Reliable identification of people is required for many applications such as immigration control, aviation security, or safeguarding of financial transactions. Research has demonstrated that the texture of a person's iris is unique and can be used as a means of identification [1]. Improving the accuracy and reliability of iris recognition is the goal of many current research endeavors [2].

The canonical iris recognition system involves a number of steps [2]. First, a camera acquires an image of an eye. Next, the iris is located within the image. The annular region of the iris is "unwrapped", or transformed from raw image coordinates to normalized polar coordinates. A texture filter is applied to numerous locations on the iris, and the filter responses are quantized to yield a binary iris code. Finally, the iris code is compared with a known iris code in the gallery, and a similarity or distance score is reported. In an identity-verification application, the system uses the reported score to decide whether the two compared iris codes are from the same subject or from different subjects.

A. Fragile Bit Masking

Not all bits in the iris code are consistent across different images of the same iris. The concept that some bits in the iris code are less consistent than others was first suggested by Bolle et al. [3]. We can improve the traditional iris recognition system by masking bits of the iris code which are less consistent [4]. The system applies Gabor filters to a number of locations on an iris image and obtains a complex filter response for each location. Each complex filter response is quantized to two bits; the first bit is set to one if the real part of the complex number is positive, and the second bit is one if the imaginary part is positive.

Consider multiple images of the same iris. A filter applied to one location on the iris produces a complex value. Across all images, the complex values for that location will be similar, but not exactly the same. Similarly, the bit from the binary quantization could be the same across all iris codes, or it may differ in some of the codes. A bit in a subject's iris code is *consistent* if it assumes the same value for most images of that subject. A bit is *fragile* if it varies in value some substantial percent of the time.

For a filter applied to a specific location in a single image, if the real part of the complex number has a large magnitude, then the corresponding bit will likely be consistent. On the other hand, if the real part is close to zero (or close to the vertical axis of the complex plane), the corresponding bit is fragile. Similar logic applies to the imaginary bits.

To illustrate this concept, we took 54 images of the same iris, and for each image, looked at the filter response for one particular spot on the iris. The resulting 54 complex numbers are shown in Figure 1. For this spot on the iris, all of the filter responses had a positive real value, but the imaginary part was positive about half of the time and negative the other half of the time. Therefore, the corresponding real bit in the iris code was consistent, and the corresponding imaginary bit was fragile.

This concept suggests a simple optimization to iris recognition algorithms. When generating an iris code, we sort the real parts of the complex numbers, and identify a fraction of numbers with the smallest magnitude. Daugman suggested selecting the lowest quartile of values [5]. The corresponding bits in the iris code are masked, or not considered when obtaining distance scores. The bits corresponding to the smallest imaginary values are likewise masked. With this modification, the final score in a comparison is based on fewer bits, but each bit used is more consistent. We call this modification *fragile bit masking*. For a more in-depth discussion of fragile bits, see our earlier work [4].

B. Motivation of Proposed Method

When using fragile bit masking [4], we mask a significant amount of information because it is not "stable". Rather than completely ignoring all of that fragile bit information, we would like to find a different way to use those bits. We know that the values (zero/one) of those bits are not stable. However, the physical locations of those bits should be stable and might be used to to improve our iris recognition performance.

Karen Hollingsworth, Kevin Bowyer, and Patrick Flynn (kholling, kwb, flynn@cse.nd.edu) are with the University of Notre Dame.

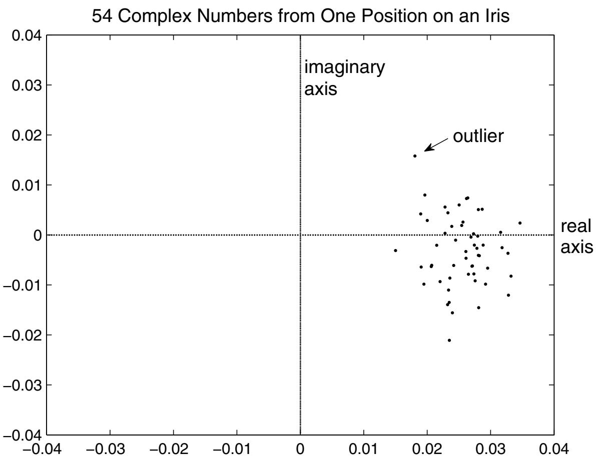


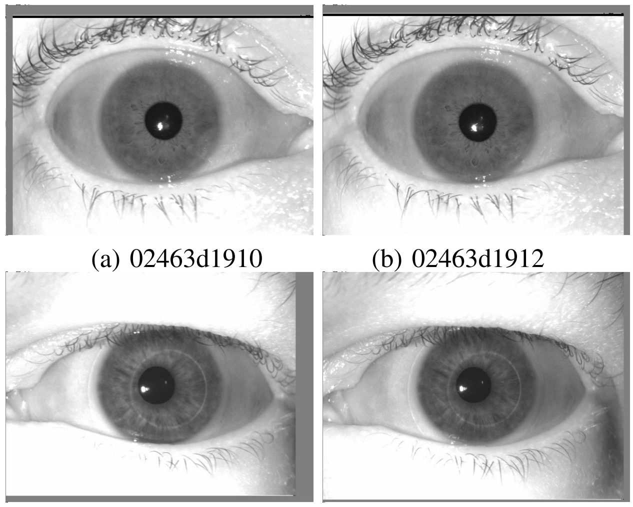
Fig. 1. These 54 complex numbers, each from the same region on 54 different images of the same subject's eye, all correspond to the same location on the iris code. Each complex number is mapped to two bits. This particular part of the iris code had a highly consistent real bit, and a highly inconsistent imaginary bit. (Figure reprinted from Hollingsworth et al., *IEEE Transactions on Pattern Analysis and Machine Intelligence* [4] ©2008 IEEE.)

We call the physical locations of fragile bits a *fragile bit pattern*. Figure 2 shows some iris images and Figure 3 shows the corresponding fragile bit patterns. Figure 3 A and B both show subject number 2463, and C and D both show subject 4261. The fragile bit patterns in A and B are similar to each other, but the fragile bit patterns in A and C are different.

To compute the Hamming distance between two iris codes, we must first combine (AND) the masks of the two iris codes. Figure 4 shows the fragility masks obtained by ANDing pairs of fragility masks together. For example, Figure 4A is the comparison mask obtained by combining Figure 3A and 3B. Figure 4A and B both show masks obtained when computing the Hamming distance for a match comparison (same subject). C and D show masks for nonmatch comparisons. The fragile bit patterns for the match comparisons line up more closely than the fragile bit patterns for the nonmatch comparisons. By looking at how well two fragile bit patterns align, we can make a soft prediction on whether those two irises are from the same subject or from different subjects. We can fuse that information with the Hamming distance information and get an improved prediction over a baseline method of using Hamming distance alone.

C. Organization of paper

The rest of this paper is organized as follows. In section II we talk about related research. Section III describes the data set used for our experiments. Section IV defines a new metric, the *fragile bit distance* (FBD), which quantifies the difference between two fragile bit patterns. In section V we talk about how to fuse Hamming distance (HD) with FBD. In section VI we show that the proposed method results in a statistically significant improvement over using Hamming distance alone.



(a) 02463d1910 (b) 02463d1912

(c) 04261d1032 (d) 04261d1034

Fig. 2. Example images from our data set. These images were captured using an LG4000 iris camera.

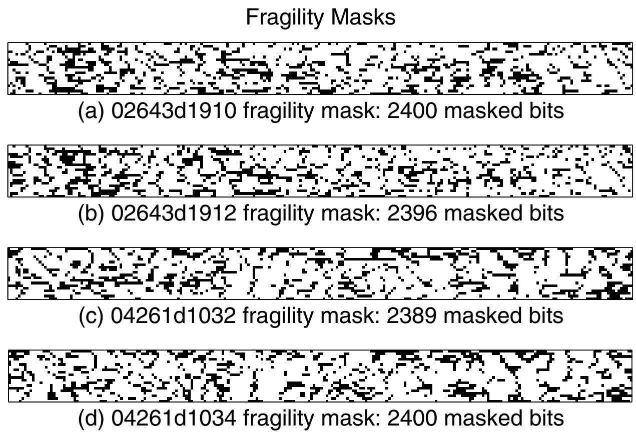


Fig. 3. These are the fragile bit patterns (imaginary part) corresponding to the images in Figure 2. Black pixels are bits masked for fragility. We use 9600-bit iris codes and mask 25% of the bits (or 2400 bits) for fragility. For partially occluded eyes, some of the bits are masked for occlusion, and so slightly less than 2400 bits are masked for fragility.

II. RELATED WORK

A. Research on fusing Hamming distance with added information

There are many papers available presenting research in iris biometrics. For a broad survey of this work, readers may look at [2]. On the other hand, the research that investigates combining Hamming distance with other information is a small subset of iris biometrics research.

Sun et al. [6] aim to characterize global iris features using the following feature extraction method. First, they introduce a *local binary pattern* operator (LBP) to characterize the iris texture in each block of the iris image. The image block information is combined to construct a global graph. Finally, the similarity between two iris images is measured using a graph matching scheme. They fuse the LBP score with

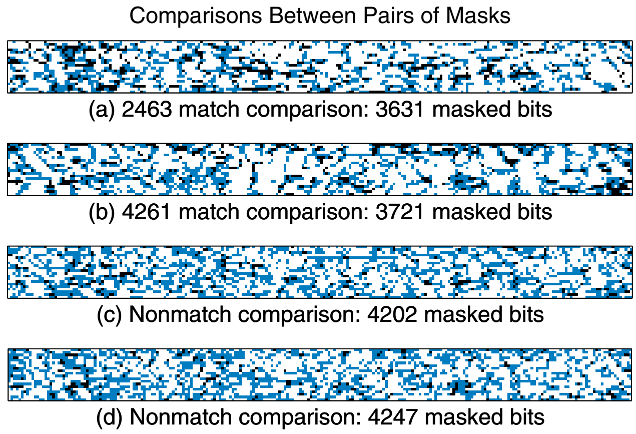


Fig. 4. These are comparisons of fragile bit patterns, each obtained by ANDing two fragile bit masks together. For example, Figure 4A is the comparison mask obtained by combining Figure 3A and 3B. Black pixels show where the two masks agreed. Blue pixels show where they disagreed. White pixels were unmasked for both iris codes. There is more agreement in same-subject comparisons than there is when comparing masks of different subjects.

Hamming distance using the Sum rule. They report that using Hamming distance alone yields an equal error rate (EER) of 0.70%, but the score-fusion of Hamming distance with their LBP method yields an EER of 0.37%.

As an alternative to the sum rule, they also note that the LBP score could be combined with Hamming distance using cascaded classifiers. Since their LBP method is slower than computing the Hamming distance, they suggest calculating the Hamming distance first. If the Hamming distance is below some low threshold, the comparison is classified as a match. If the Hamming distance is above some high threshold, the comparison is classified as a nonmatch. If the Hamming distance is between those two thresholds, use the second classifier (the LBP) to make the decision.

Vatsa et al. [7] characterize iris texture using Euler numbers. They use a Vector Difference Matching algorithm to compare Euler codes from two irises. Vatsa et al. combine Hamming distance and Euler score using a cascaded classifier.

Zhang et al. [8] use log Gabor filters to extract 32 global features characterizing iris texture. To compare the global features from two iris images, they use a weighted Euclidean distance (WED) between feature vectors. Zhang et al. use cascaded classifiers to combine the global WED with a Hamming distance score. However, unlike Sun et al. [6] and Vatsa et al. [7], they propose using their global classifier first, and then using Hamming distance. In their experiments, using Hamming distance alone gave a false accept rate (FAR) of 8.1% when the false reject rate (FRR) was 6.1%. The fusion of WED and Hamming distance gave FAR = 0.3%, FRR = 1.9%.

Park and Lee [9] generate one feature vector using the binarized directional subband outputs at various scales. To compare two binary feature vectors, they use Hamming distance. A second feature vector is computed as the blockwise normalized directional energy values. Energy feature vectors are compared using Euclidean distance. To combine scores from these two feature vectors, Park and Lee use a weighted average of the two scores. Using the binary feature vectors alone gives an EER of 5.45%; the energy vectors yield an EER of 3.80%; when the two scores are combined, the EER drops to 2.60%.

All of the above mentioned papers combine Hamming distance scores with some other scores at the matching score level to improve iris recognition. Sun et al. [6] combine scores by summing. Three of the papers [6], [7], [9] use cascaded classifiers. Park and Lee [9] use a weighted average. Our work is similar to these papers in that we also consider combining two match scores to improve performance. We differ from these other works in that we are the first to use a score based on the location of fragile bits in two iris codes.

B. Research on Fragile Bits

Research on fragile bits is a more recent trend in iris biometrics literature. One of our previous papers, [4], presented evidence that not all bits in the iris code are of equal consistency. We investigated the effect of different filters on bit fragility. We used 1D log-Gabor filters and multiple sizes of a 2D Gabor filter, and found that the fragile bit phenomenon was apparent with each filter tested. The largest filter tended to yield fewer fragile bits than the smaller filters. We investigated possible causes of inconsistencies and concluded that the inconsistencies are largely due to the coarse quantization of the filter response. We performed an experiment comparing (1) no masking of fragile bits (baseline) with (2) masking bits corresponding to complex filter responses near the axes of the complex plane. We masked fragile bits corresponding to the 25% of filter responses closest to the axes. Using a data set of 1226 images from 24 subjects, we found that fragile bit masking improved the separation between the match and nonmatch score distributions.

Other researchers have also begun to investigate the effects of masking fragile bits. Barzegar et al. [10] investigated fragile bit masking using different thresholds. They compared (1) no fragile bit masking to (2) fragile bit masking with thresholds of 20%, 30% and 35%. They found that using a threshold of 35% for masking produced the lowest error rates on the CASIA-IrisV3 data set. Our own initial investigations have shown that the optimal fragility threshold may depend partly on the quality of the iris images being used; therefore, we feel that further investigation into the proper fragility threshold would be worthwhile.

Dozier et al. [11] also tried masking inconsistent bits and found an improvement in performance. However, they used a different method than Hollingsworth et al. [4] and Barzegar et al. [10]. Hollingsworth et al. [4] and Barzegar et al. [10] approximated fragile bit masking by masking filter responses near the axes of the complex plane. In contrast, Dozier et al. used a training set of ten images per subject to find consistency values for each bit in the iris code. Then for that subject, they only kept bits that were 90% or 100%

Fig. 5. Images in our data set were captured using this LG4000 iris camera [13].

consistent in their training set, and masked all other bits. In addition, they also considered only those bits that had at least 70% coverage in their training set; that is, if a bit was occluded by eyelids or eyelashes in four or more of the training images, they masked that bit. Dozier et al. tested their method on six subjects from the ICE data set.

In a similar paper, Dozier et al. [12] again showed the benefit of masking inconsistent bits. In this work, they used a genetic algorithm to create an iris code and corresponding mask for each subject. Once again, they used ten training images per subject in generating their fragile bit masks.

Each of the above mentioned papers showed the benefit of masking fragile bits, but in every case, they simply discarded all information from the fragile bits. None of them considered employing the locations of fragile bits as an extra feature to fuse with Hamming distance.

III. DATA

We acquired a data set of 9784 iris images taken with an LG4000 iris camera [13] at the University of Notre Dame. Some example images are shown in Figure 2 and the camera is shown in Figure 5. All images in this set were acquired in 2008. A total of 460 different people attended acquisitions sessions, so there are 920 different eyes in the data set. Each subject attended between one and twelve acquisition sessions throughout the year. At each session, we acquired three left eye images and three right eye images. The minimum number of images for any one subject is six (three of each iris), and the maximum is 72.

IV. FRAGILE BIT DISTANCE (FBD)

Figure 4 provides some indication of what we should expect when comparing two fragile bit patterns. In a genuine comparison, the locations of the fragile bits line up. In an impostor comparison the locations of the fragile bits do not. When we compare two iris codes, we mask any bit that is fragile in either of the two fragile bit patterns. Therefore, we expect more bits to be masked for fragility in impostor comparisons than in genuine comparisons.

We can theoretically predict how many bits will be unmasked in an impostor comparison. In this analysis, we make the assumption that the fragility of bits is independent of position and that each position is independent of all other positions. Consider the iris code for a single, unoccluded image. We mask 25% of bits for fragility, and leave 75% of bits unmasked. Now consider a comparison of two unoccluded images from different subjects. We expect (75%)(75%) = 56.25% of the bits to be unmasked, and 43.75% of the bits to be masked.

Another way to analyze how many bits will be masked is to consider the number of coincident bits. If we mask 25% of the bits in each of the two irises in an impostor comparison, we expect (25%)(25%) = 6.25% of the bits to be coincident fragile bits. About $25\% - 6.25\% = 18.75\%$ of the bits in the first iris code will be marked as fragile and *not* line up with any fragile bits from the second iris code. The total number of masked bits *for the comparison* will be the coincident fragile bits, plus the bits masked in first iris code only, plus the bits masked in the second iris code only. Therefore we expect 6.25% + 18.75% + 18.75% = 43.75% of the unoccluded bits will be masked in an impostor comparison.

In contrast, a genuine comparison will have fewer masked bits. In two identical images, the fragile bits will line up exactly and the comparison will have 75% unmasked bits and 25% masked bits. However, two different images of the same iris are not identical because of differences in lighting, dilation, distance to the camera, focus, or occlusion. Therefore, on average, more than 25% of the bits will be masked in a genuine comparison.

We define a metric called the *fragile bit distance* (FBD) to measure how well two fragile bit patterns align. In order to compute fragile bit distance, we need to store occluded bits and fragile bits separately. Therefore, each iris template will consist of three matrices: an iris code i, an occlusion mask m, and a fragility mask f. Unmasked bits are represented with ones and masked bits are represented with zeros. Specifically, unoccluded bits and consistent bits are marked as ones, while occluded and fragile bits are zeros. We do not want FBD to be affected by occlusion, so we consider only unoccluded bits when computing the FBD.

Take two iris templates, template A and template B. The FBD is computed as follows:

$$FBD = \frac{\|m_A \cap m_B \cap \overline{f_A \cap f_B}\|}{\|m_A \cap m_B\|} \tag{1}$$

where $\cap$ represents the AND operator, and the line over $\overline{f_A \cap f_B}$ represents the NOT operator. The norm ($\|\|$) of a matrix tallies the number of ones in the matrix.

In above equation, $\overline{f_A \cap f_B}$ is a matrix storing all bits masked for fragility. $m_A \cap m_B$ is a matrix marking all bits unoccluded by eyelashes and eyelids. The FBD expresses the fraction of unoccluded bits masked for fragility in the comparison. This metric is large for impostor comparisons, and small for genuine comparisons.

Our theory predicts that we will have an average FBD of 0.4375 for impostor comparisons, and an average FBD of somewhere between 0.25 and 0.4375 for genuine comparisons. We tested these predictions on our data set of 9784 images. The average FBD for genuine and impostor comparisons are reported in Table I, with standard deviations reported in parentheses.

Table I: Average FBD for Genuine and Impostor Comparisons (std dev)

	Avg. Genuine FBD	Avg. Impostor FBD
Theoretical value	between 0.25 & 0.4375	0.4375
LG4000 images	0.4058 (0.0142)	0.4393 (0.0082)

The average impostor FBD was within one standard deviation of the theoretical prediction. Also, the average genuine FBD was less than the average impostor FBD.

V. Fusing Fragile Bit Distance with Hamming Distance

The fragile bit distance (FBD) quantifies the difference between two fragile bit masks. A comparison with a low FBD is likely to be a comparison between two images of the same subject and a comparison with a high FBD is likely to be a comparison between two different subjects. However, when used alone, the FBD is a worse classifier than the Hamming distance (HD) metric. Figure 6 shows how the FBD relates to the Hamming distance for some comparisons. Each point in this figure represents one comparison between two iris codes. Impostor comparisons are graphed in red and genuine comparisons are graphed in blue. Selecting a single threshold of Hamming distance (e.g. HD = 0.35) would separate genuine and impostor comparisons better than any threshold we might choose for FBD.

Even though the FBD is not as powerful a metric as the Hamming distance, we can combine the features to create a better classifier than Hamming distance alone. When using Hamming distance alone, we achieve an equal error rate of 9.227×10^{-3}. To combine Hamming distance and FBD, we first tried a weighted average technique, using the same approach as [9]. We combined the two scores using the equation,

$$Score_W = \alpha \times HD + (1 - \alpha) \times FBD. \tag{2}$$

We varied the parameter α in steps of 0.1 from 0 to 1, and calculated the equal error rate for each run. Figure 7 shows how the equal error rate changes as α varies. The lowest equal error rate was 8.455×10^{-3}, which was obtained using an α value of 0.4. However, the difference between $\alpha = 0.4$ and $\alpha = 0.5$ was insignificant.

Multiplication can be used as an alternative method of score fusion:

$$Score_M = HD \times FBD. \tag{3}$$

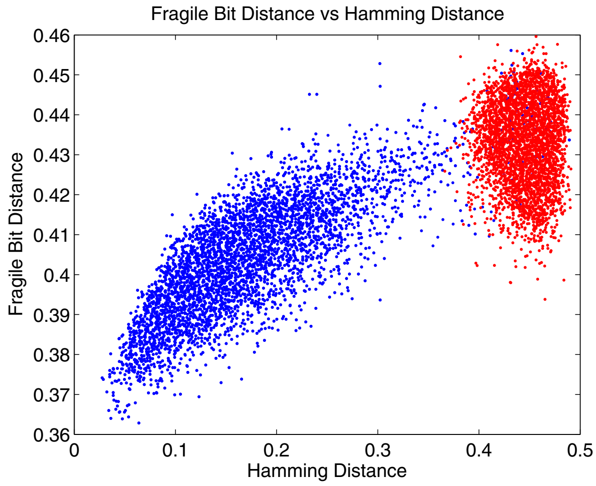


Fig. 6. Each dot in this graph represents a comparison between two iris codes. The Hamming distance and FBD for each comparison are shown. Blue dots represent genuine comparisons, and red dots represent impostor comparisons. This graph shows that FBD has less predictive power than the Hamming distance. However, when both metrics are combined, the resulting classifier performs better than using Hamming distance alone.

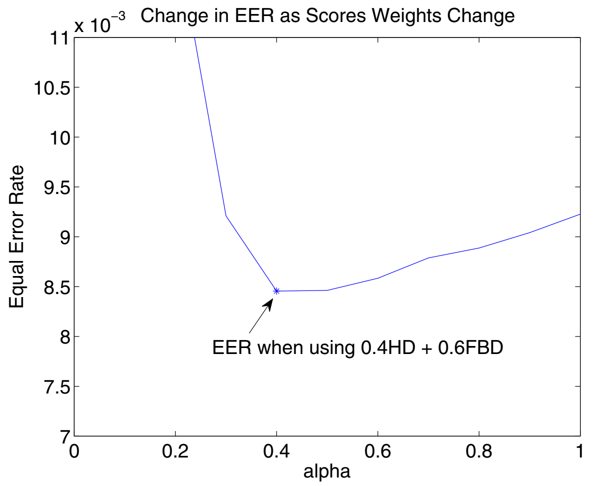


Fig. 7. We fused FBD and HD using the expression, $\alpha \times HD + (1 - \alpha) \times FBD$. We found that an α value of 0.4 yielded the lowest equal error rate.

When using multiplication, the equal error rate was 8.451×10^{-3}. Fusing by multiplication and fusing by weighted average yielded nearly identical results. An ROC curve showing the results of these tests is shown in Figure 8, and Table II shows summary statistics of these experiments including the equal error rate (EER), the false reject rate at an operating point of FAR=0.001 (FRR at FAR=0.001), and the area under the ROC curve (AUROC).

Based on the values in Table II, we see that both methods of fusing Hamming distance with FBD performed better than using Hamming distance alone. By incorporating FBD, we improved the accuracy of our iris matcher. The difference between using a weighted average and multiplying was insignificant. Since the EER for the multiplication method was slightly smaller, we chose to use that method of fusing

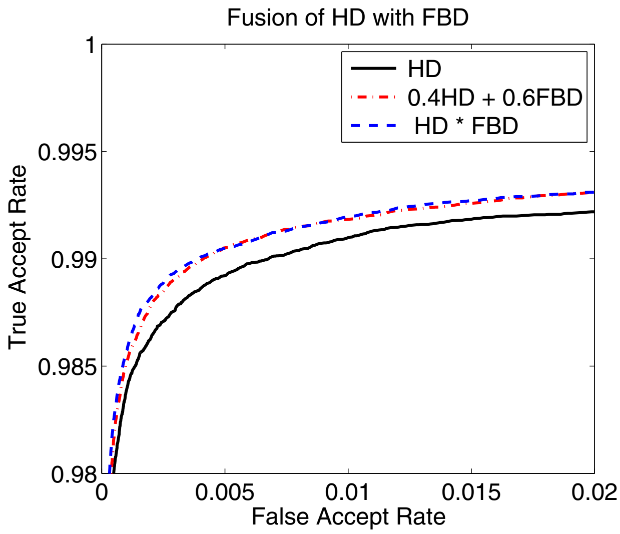


Fig. 8. Fusing Hamming distance with FBD performs better than using Hamming distance alone. Fusing by multiplying and fusing by weighted averaging yield similar results.

Table II: Fusing FBD with Hamming distance

Method	EER	FRR at FAR=0.001	AUROC
HD (baseline)	9.23×10^{-3}	1.62×10^{-2}	0.9976
0.4HD + 0.6FBD	8.45×10^{-3}	1.49×10^{-2}	0.9977
HD $\times$ FBD	8.45×10^{-3}	1.43×10^{-2}	0.9978

Hamming distance with FBD.

One caveat with using FBD is that in order to compute FBD, we have to store the fragility mask separately from the occlusion mask. Therefore, our iris template is slightly bigger (1.5 times bigger) than it would be if we did not use FBD.

VI. TEST OF STATISTICAL SIGNIFICANCE

The proposed fusion between Hamming distance and FBD works better than the baseline test of Hamming distance alone. The equal error rate (EER) for the baseline is 9.23×10^{-3} and the EER for the proposed fusion of Hamming distance and FBD is 8.45×10^{-3}. To test for statistical significance, we randomly divided the 460 subjects into ten different test sets. For each test set, we measured the performance of using Hamming distance alone, and of using fusion of Hamming distance and FBD. Then we used a paired t-test to see whether the proposed method obtained a statistically significant improvement. The results are given in Table III. The t-test showed statistically significant improvement of the proposed method over the baseline for both EER and false reject rate at a false accept rate of 0.1% (FRR at FAR=0.001). Rerunning the same experiment using different random test sets gave very similar results.

VII. CONCLUSION

In this paper, we defined a new metric, the *fragile bit distance* (FBD) which measures how two fragile bit masks differ. Low FBDs are associated with genuine comparisons

Table III: Is HD $\times$ FBD better than HD alone?

Method	Avg. EER	Avg. FRR at FAR=0.001
HD (baseline)	9.40×10^{-3}	1.71×10^{-2}
HD $\times$ FBD (proposed)	8.55×10^{-3}	1.48×10^{-2}
p-value	6.6×10^{-3}	3.46×10^{-2}

between two iris codes. High FBD are associated with impostor comparisons.

The multiplication of FBD and Hamming distance is a better classifier than using Hamming distance alone. By multiplying FBD and Hamming distance, we reduced the EER of our iris recognition system from 9.23×10^{-3} to 8.45×10^{-3}, a statistically significant improvement.

ACKNOWLEDGEMENT

This research is supported by the National Science Foundation under grant CNS01-30839, by the Central Intelligence Agency, by the Intelligence Advanced Research Projects Activity, and by the Technical Support Working Group under US Army contract W91CRB-08-C-0093. 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.

REFERENCES

[1] John Daugman. How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):21–30, 2004.

[2] Kevin W. Bowyer, Karen P. Hollingsworth, and Patrick J. Flynn. Image understanding for iris biometrics: A survey. *Computer Vision and Image Understanding*, 110(2):281–307, 2008.

[3] Ruud M. Bolle, Sharath Pankanti, Jonathon H. Connell, and Nalini Ratha. Iris individuality: A partial iris model. In *Proc. Int. Conf. on Pattern Recognition*, pages II: 927–930, 2004.

[4] Karen P. Hollingsworth, Kevin W. Bowyer, and Patrick J. Flynn. The best bits in an iris code. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(6):964–973, Jun 2009.

[5] John Daugman. personal communication.

[6] Zhenan Sun, Tieniu Tan, and Xianchao Qiu. Graph matching iris image blocks with local binary pattern. In *Proc. Int. Conf. on Biometrics (Springer LNCS 3832)*, pages 366–372, Jan 2006.

[7] Mayank Vatsa, Richa Singh, and Afzel Noore. Reducing the false rejection rate of iris recognition using textural and topological features. *Int. Journal of Signal Processing*, 2(2):66–72, 2005.

[8] Peng-Fei Zhang, De-Sheng Li, and Qi Wang. A novel iris recognition method based on feature fusion. In *Proc. Int. Conf. on Machine Learning and Cybernetics*, pages 3661–3665, 2004.

[9] Chul-Hyun Park and Joon-Jae Lee. Extracting and combining multimodal directional iris features. In *Int. Conf. on Biometrics (Springer LNCS 3832)*, pages 389–396, Jan 2006.

[10] Nakissa Barzegar and M. Shahram Moin. A new user dependent iris recognition system based on an area preserving pointwise level set segmentation approach. *EURASIP Journal on Advances in Signal Processing*, 2009.

[11] Gerry Dozier, David Bell, Leon Barnes, and Kelvin Bryant. Refining iris templates via weighted bit consistency. In *Proc. Midwest Artificial Intelligence and Cognitive Science (MAICS) Conference*, Apr 2009.

[12] Gerry Dozier, Kurt Frederiksen, Robert Meeks, Marios Savvides, Kelvin Bryant, Darlene Hopes, and Taihei Munemoto. Minimizing the number of bits needed for iris recognition via bit inconsistency and grit. In *Proc. IEEE Workshop on Computational Intelligence in Biometrics: Theory, Algorithms, and Applications*, Apr 2009.

[13] LG IrisAccess 4000. http://www.lgiris.com/ps/products/irisaccess4000.htm, accessed Apr 2009.