Information Theory and the IrisCode

John Daugman

***Abstract*— Iris recognition has legendary resistance to false matches, and the tools of information theory can help to explain why. The concept of entropy is fundamental to understanding biometric collision avoidance. This paper analyses the bit sequences of IrisCodes computed both from real iris images and from synthetic white noise iris images, whose pixel values are random and uncorrelated. The capacity of the IrisCode as a channel is found to be 0.566 bits per bit encoded, of which 0.469 bits of entropy per bit is encoded from natural iris images. The difference between these two rates reflects the existence of anatomical correlations within a natural iris, and the remaining gap from one full bit of entropy per bit encoded reflects the correlations in both phase and amplitude introduced by the Gabor wavelets underlying the IrisCode. A simple two-state hidden Markov model is shown to emulate exactly the statistics of bit sequences generated both from natural and white noise iris images, including their imposter distributions, and may be useful for generating large synthetic IrisCode databases.**

***Index Terms*— Entropy, IrisCode, hidden Markov models.**

I. Introduction

INFORMATION theory [1] analyses relationships between random variables using metrics that quantify what they convey probabilistically about each other. Its methods are suited for domains such as inference, communication channels, classifiers and pattern recognition generally, but it has been little used in the field of biometrics except in connection with cryptographic protocols. This is odd, because many concepts in biometrics correspond closely with the idea of a *noisy channel*, its *capacity* or that of an encoding scheme, as well as the *entropy* of a random variable or code. For example, the randomness and complexity of biometric patterns are generally understood to determine their uniqueness and hence their ability to avoid collisions with others (False Matches), but these biometric properties are rarely quantified in terms of entropy. Now that (for example) almost a billion persons have had their IrisCodes enrolled in a national ID deployment across India [2], [3], each being compared with all others for de-duplication checks, it is time that the origin of biometric collision avoidance become more quantitatively and widely understood. The canonical Venn diagram of Fig. 1 summarises some key information-theoretic concepts involving uncertainty, various entropies defined by probability distributions, and mutual information between random variables X and Y.

Manuscript received May 1, 2015; revised August 9, 2015; accepted October 22, 2015. Date of publication November 12, 2015; date of current version December 10, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Venu Govindaraju.

The author is with the Faculty of Computer Science and Technology, University of Cambridge, Cambridge CB3 0FD, U.K. (e-mail: john.daugman@cl.cam.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2015.2500196

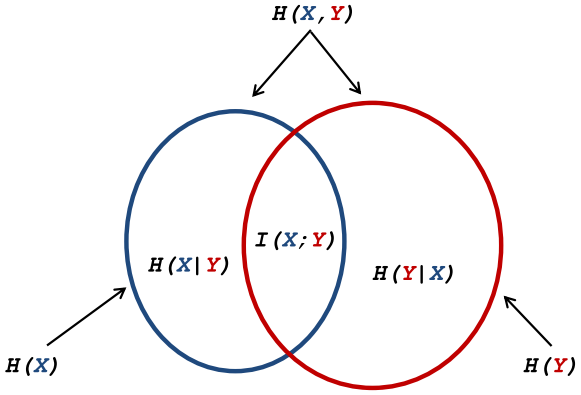

Fig. 1. Classical Venn diagram of relationships between random variables, their entropies and mutual information. Applied to biometrics, X and Y might be identities, images, biometric features, computed templates, or decisions.

For purposes of illustration, we might take random variable X to represent identity of persons, and Y to represent biometric signals. There is uncertainty about both, represented by the overlapping ovals labelled $H(X)$ and $H(Y)$. If a discrete random variable (say Y) has n possible states, and the i^{th} of these occurs with probability p_i, the information gained by observing that state is defined to be $\log_2 p_i$ bits. To gain information is to lose uncertainty (entropy) by the same amount, so entropy is defined as the negative of information and hence is non-negative. One reason for the logarithmic measure is to cause the information gained from observing independent events to be additive, since the joint probability of independent events (i, j) simply multiplies $(p_i p_j)$ and thus the information gained from their joint observation, the log of this product, is just the sum of the information gained from the individual observations. Summing contributions over the entire ensemble of n possible states, weighting each by its probability p_i of occurence and with $\sum_i p_i = 1$, we get the classic Shannon measure [1] of the *entropy* (in bits) of this random variable:

$$H(Y) = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{1}$$

$H(Y)$ increases with the number n of possible states, and for any given n it is maximised at $\log_2 n$ if all the possible states are equiprobable: $\forall i,\ p_i = 1/n$. The information capacity of a coding scheme is characterised partly by its entropy H in bits, because 2^H is an upper limit on the number of distinct states that can be encoded in the worst case that they are equiprobable. Larger entropy H means that more objects can be uniquely coded (collision avoidance, *i.e.* distinctiveness), but the entropy of a code is reduced by non-independence of its bits, and the entropy of a population is reduced by

non-uniformity in its probability distribution. There is a strong analogy with cryptography: the strength of a cryptographic key grows with its entropy (maximally its length in bits), but it is weakened by any non-randomness or predictability, as famously helped British code-breakers at Bletchley Park in WWII routinely break German Enigma codes.

The joint entropy $H(X, Y)$ of random variables X and Y, as well as their conditional entropies $H(X|Y)$ and $H(Y|X)$ as demarcated in Fig. 1, are defined in the same way as (1) except that the probabilities used are instead the joint or conditional probabilities: $p(x, y)$, $p(x|y)$, and $p(y|x)$ where x and y are the values that may be taken by the random variables X and Y. In the present biometric context the crescent-shaped region on the left, the conditional entropy $H(X|Y)$, would be interpreted as: *"how much uncertainty remains about personal identity X, given the biometric measurements Y."* Likewise the inverse conditional entropy $H(Y|X)$, the crescent-shaped region on the right, would be interpreted as: *"given identity X, how much uncertainty remains about the biometric signals Y that it may generate."* Obviously both of these conditional entropies should be minimised, as they are the origins of False Matches and False non-Matches.

The area of intersection between the ovals in Fig. 1 is the mutual information, $I(X; Y)$. The larger it is, the better, since it signifies how much these two random variables convey about each other. Ideally, biometric patterns Y should leave no uncertainty about identities X, and likewise, identities X should generate consistent and stable biometric patterns Y. Departures from these ideal mappings are reflected in reductions in the mutual information. In the worst case ("biometric uselessness"), the two ovals are disjoint and non-overlapping: personal identity, and biometric patterns, say nothing about each other and are independent. In the best case ("biometric determinism"), $H(X)$ and $H(Y)$ are fully co-extensive with each other and with their mutual information $I(X; Y)$, and the conditional entropies $H(X|Y)$ and $H(Y|X)$ are both nil.

We come finally to the key idea of a *channel*, which is relevant in many different biometric contexts. It has an input (random variable X) and an output (random variable Y). The fidelity between X and Y is measured in terms of their mutual information $I(X; Y)$, which leads to one measure of *channel capacity*. We could regard X and Y as identities, or as biometric patterns, either or both. The fraught process of biometric presentation, with inconsistent acquisition, could itself be regarded as an internal part of the channel. For illustration initially let us consider a channel for example just as a biometric image coding scheme (see Fig.s 2 and 3).

The goal of this coding channel is to represent accurately at the output what is actually present at the input. Iris images are encoded using their projections onto 2D Gabor wavelets, which can extract all their information in a very compact representational format. In Fig. 3, the upper panels show two original iris images from different ethnic groups, and the lower panels show their corresponding reconstructions by linear combinations of 2D Gabor wavelets drawn from a self-similar family of wavelets having six discrete orientations, two quadrature phases, and five sizes (or frequencies) in an octave scaling sequence spanning four octaves. This discrete

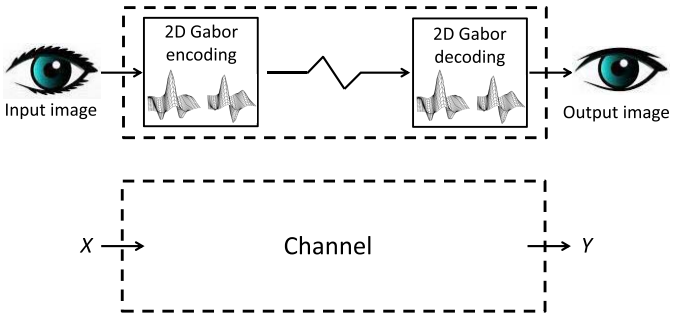

Fig. 2. The concept of a *channel* between input and output random variables X and Y is central in information theory and it offers quantitative metrics applicable to many mappings used in biometrics, whether between identities and templates, or simply an image coding scheme.

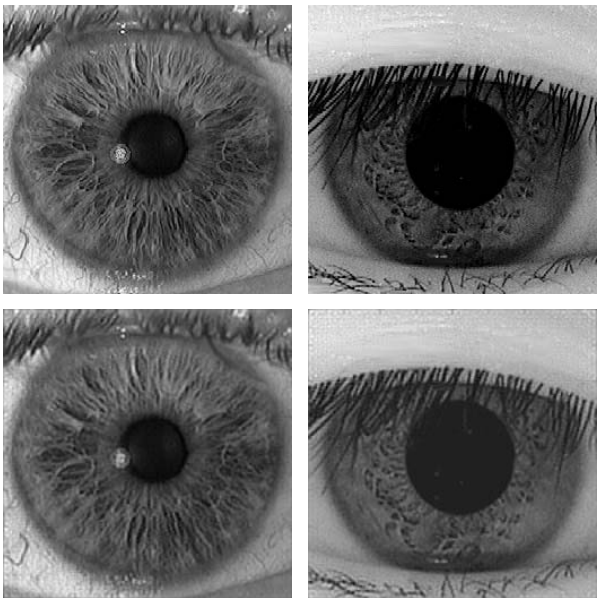

Fig. 3. Demonstration that the image analysis scheme underlying the IrisCode can also constitute a complete image code. The upper panels show two original iris images, while the lower panels show their reconstruction simply by adding together 2D Gabor wavelets, in a discrete array of orientations, sizes, and positions, with appropriate coefficients that constitute the code.

set of wavelets on a sparse spatial lattice has a highest spatial frequency that is lower than the pixel resolution, and so the reconstruction is imperfect (as close inspection reveals). Image quality metrics such as mean-squared-error or signal-to-noise ratio would characterise this coding scheme in terms of mutual information and channel capacity. But this paper will focus instead on encodings for automatic biometric identification, using the same class of wavelets that generated the image coding in Fig. 3.

II. IrisCode Properties

All public deployments of iris recognition are based on the IrisCode, although several alternative variants have been proposed and studied in the academic literature. The IrisCode and its match engine using a Hamming Distance (HD) metric have been extensively discussed already ([4]–[8]) and will be only briefly summarised here. After localisation, segmentation, and normalisation of the iris tissue, its random texture is

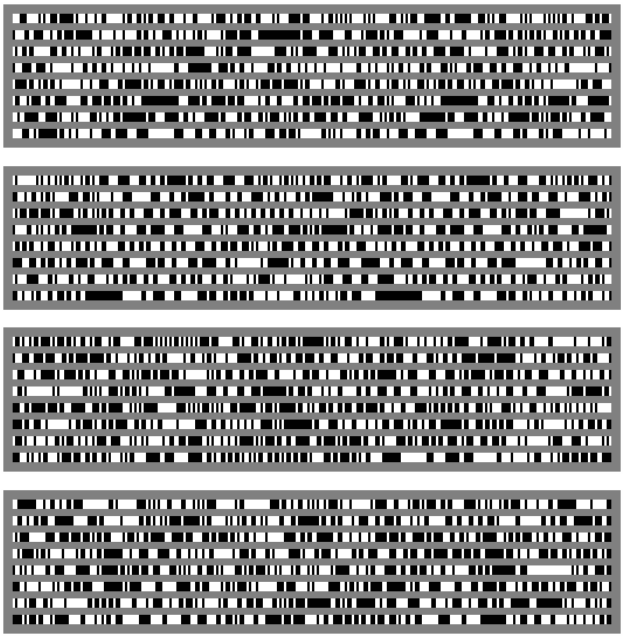

Fig. 4. Graphical portrayal of the IrisCodes produced by four different eyes. In each case, the bit streams from two different wavelets are concatenated.

encoded into a bit stream using the sign of its local projections onto a parameterised family of 2D Gabor wavelets:

$$h_{\{\mathrm{Re,Im}\}} = \mathrm{sgn}_{\{\mathrm{Re,Im}\}} \int_{\rho}\int_{\phi} I(\rho,\phi) e^{-i\omega(\theta_0-\phi)} e^{-(r_0-\rho)^2/\gamma^2} e^{-(\theta_0-\phi)^2/\beta^2} \rho d\rho d\phi \tag{2}$$

where $h_{\{\mathrm{Re,Im}\}}$ can be regarded as a complex-valued bit whose real and imaginary parts are either 1 or 0 (signum) depending on the sign of the 2D integral; $I(\rho,\phi)$ is the iris image normalised into a dimensionless pseudo-polar coordinate system; γ and β are the size parameters of multi-scale 2D Gabor wavelets, spanning an 8-fold range corresponding to 0.15mm to 1.2mm on the iris; ω is wavelet frequency, spanning three octaves in inverse proportion to β such that the profiles are roughly as shown in the wiremesh plots within Fig. 2; and (r_0, θ_0) represent the central polar coordinates of each local patch of iris tissue for which such IrisCode bits are computed. Gabor wavelets were chosen because they optimise information resolution simultaneously in space (location) and in frequency: they have minimal joint uncertainty under the Heisenberg Uncertainty Principle. Altogether 2,048 such bits (256 bytes) are computed for encoding each iris pattern; but in addition an equal number of masking bits are also computed to signify whether any iris region is obscured by eyelids, contains any eyelash occlusions, specular reflections, boundary artifacts of hard contact lenses, or poor signal-to-noise ratio (the lowest quartile by amplitude), and thus should be ignored as artifact. This deployed mechanism for masking unreliable bits was first disclosed 15 years ago [9], but in some more recent literature it has been renamed "fragile bit" masking [10], [11].

The IrisCode bit sequences portrayed pictorially in Fig. 4 from four different eyes immediately convey the impression of large entropy, the basis for the IrisCode's legendary resistance to False Matches (collision avoidance). For example, the Government of India is now two-thirds finished with its gargantuan project to enroll the IrisCodes of all 1.2 billion citizens within three years, requiring a million enrollments per day across about 36,000 enrollment stations [2], [3]. Several different camera designs are deployed, which license the IrisCode algorithm to encode iris patterns for subsequent cross-matching. Most impressively, each new enrollee is then compared with all existing enrollees (more than 800 million persons now) in de-duplication checks, since incentives exist to try to acquire multiple identities and thereby gain fraudulent multiple access to benefits and entitlements [3]. Thus, some 800 trillion (8×10^{14}) cross-comparisons are performed every day. This requires both the great speed of Exclusive-OR (XOR) IrisCode matching, which executes at millions/sec per CPU single core, but even more importantly, the large entropy in IrisCodes [5], [6] to avoid False Matches. Weaker biometrics such as face recognition would utterly drown in False Matches at this scale, given the vast number of opportunities for biometric collisions.

A. Losses in IrisCode Entropy

The bits comprising an IrisCode are far from independent. One reason is because iris patterns contain internal correlations of natural structure, such as radial furrows that may extend all the way from the pupil to the limbus. In the orthogonal (angular) direction, many features subtend a significant angle around the pupil, as could be seen in Fig. 3, again creating spatial correlations. But an even greater source of correlations (and hence loss of IrisCode entropy) is the nature of the Gabor wavelet encoders themselves. Their functional form in (2) makes them bandpass filters, having both a lowpass aspect and a highpass aspect. Regardless of the input, convolution with such filters produces outputs in which neighbouring points are correlated in amplitude because of the lowpass aspect, but also more distant points are correlated in phase because of the highpass aspect. This latter point becomes more intuitive if one imagines that the bandpass region (between the highpass and lowpass filter characteristics) is quite narrowband when $\omega \gg 1/\beta$ in (2). Then the filter outputs are almost pure sinusoids of frequency ω regardless of the input, and hence there is oscillation with phase coherence that persists over a long interval. Indeed both the amplitude correlation caused by the lowpass aspect, and this phase coherence, each have an interval of persistence that is reciprocal to the spectral bandwidth of the wavelets used. Later we shall quantify this effect on entropy using "white noise" iris images whose input pixels are random and uncorrelated.

For different images of same eyes, Fig. 5 shows this phase coherence effect using two different wavelet frequencies. The IrisCodes extracted from different sets of same-eye images were scrolled relative to each other in steps around the optimal orientation that yields the best match (lowest HD). The unit of relative shift is $360°/256 \approx \pm 1.41°$ rotation. Whereas the best match obviously occurs when the IrisCodes are optimally aligned (0 shift), we see that same-eye IrisCodes can become, in a sense, "out of phase with each other" (*e.g.* at a relative shift of about ± 6 for the lower frequency wavelet). This effect

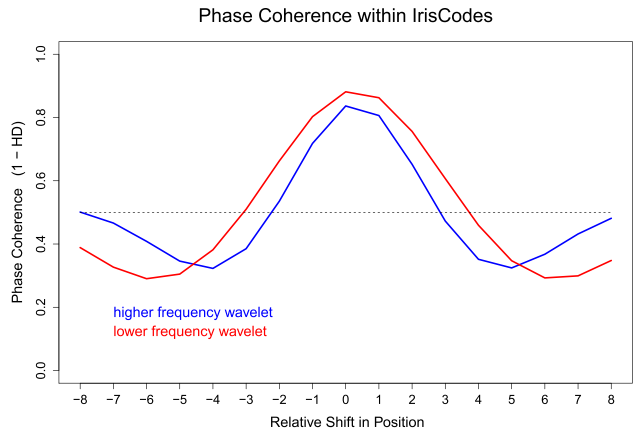

Fig. 5. Phase coherence effect in same-eye IrisCodes under relative rotation. The undulatory variation in HD scores arises because the bandpass Gabor wavelet encoders impart oscillatory phase correlations, losing some entropy.

was already observed by Matey *et al.* [12] for barrel shifts of a single IrisCode, and it was also partially observed by Rathgeb *et al.* [13] as increased variability for same-eye IrisCode comparisons under circular shifts. The biphasic coherence profiles seen in Fig. 5 occur when same-eye IrisCodes are compared, but not those from different eyes. This observation could be incorporated into the matching process as rather more "distributed" evidence of identity than using only a single best HD measurement. A somewhat paradoxical consequence of the biphasic profiles seen in Fig. 5 is that observing an unusually large fraction of *mismatching* bits (the two troughs, where 70% of the bits disagree) is actually strong evidence of a match, because only same-eye IrisCodes can get coherently "out-of-phase" with each other. This is another reflection of the key concept that iris recognition is based on the *failure* of a test of statistical independence [4].

B. Redundancy of the Real and Imaginary Parts

A consequence of phase coherence within IrisCodes is that the Re and Im parts of (2) become redundant with each other, because their quadrature phase relationship allows each to predict the other's value after a shift. The upper panel of Fig. 6 shows that without any relative shift, the IrisCode bit streams reveal no correlation between their Re and Im parts: the quadrature phase relationship of the wavelet parts makes them orthogonal, with inner product 0, and so the HD scores between them remain very close to 0.5 as expected. But under a shift corresponding to $\pi/2$ in phase (lower panel) for a given wavelet, there emerges a large correlation in the bit streams. This is not surprising, because the Re part of a Gabor wavelet when shifted in position by $\pm\pi/2$ in phase terms acquires quite a large inner product with its corresponding Im part, so they are doing similar work. A fundamental principle of information theory, which arises in several different forms in this paper, is that predictability reduces entropy. For this reason, for more than a decade all public deployments of iris recognition have used only the Re part, as will this analysis, ignoring the Im part since it adds so little further entropy.

We can see in Fig. 7 the entropy-reducing effects of bandpass encoding and quantisation when actual iris image

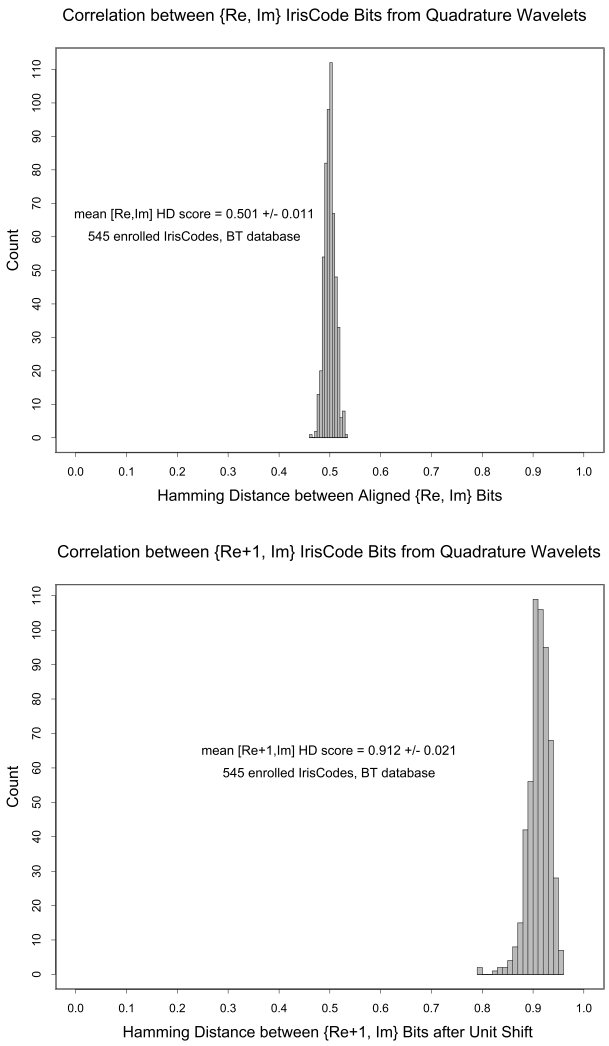

Fig. 6. A consequence of the phase coherence in IrisCodes is that the bits corresponding to the Re and Im parts cease to be independent (upper panel), and are strongly correlated under a shift of $\pm\pi/2$ in phase (lower panel).

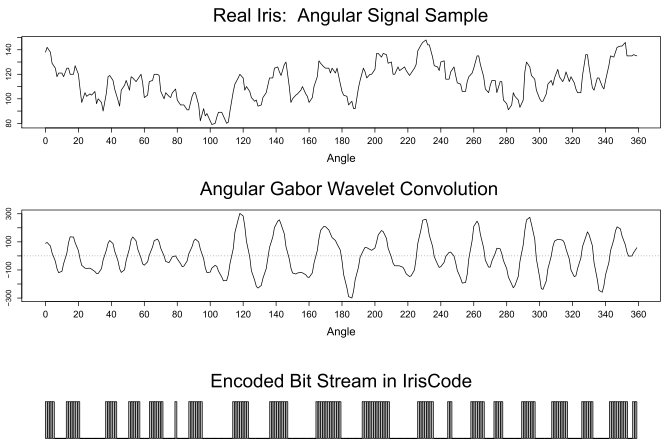

Fig. 7. Pixel values extracted from a natural iris along a single mid-radius circumference (upper trace). Gabor wavelet encoders discard slow gradients and also higher frequency structure (middle trace). The bit stream that results from quantisation (lowest trace) has properties of a "sticky oscillator."

data is passed through (2) to generate an IrisCode bit stream. The upper trace plots actual pixel values sampled from an iris along a single (mid-radius) circumference. It contains a

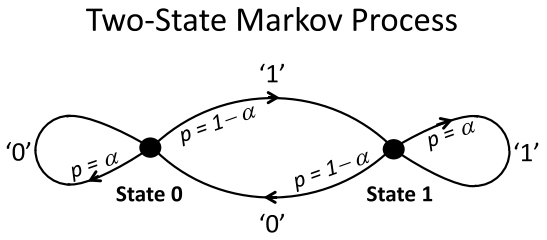

Fig. 8. Hidden Markov Model for generating IrisCode bit streams.

broad range of frequencies, including long-range trends caused by illumination gradients due to the geometry of illumination (typically from below the eye) and also from secondary light reflected from the nose onto the nasal side of the iris. Both those low frequency gradients, and also high frequency structure, are removed by the Gabor wavelet encoder (in this case a fairly low frequency cosine-phase wavelet), as shown by its output in the second trace. Finally, the signum quantisation of (2) into a bit stream is shown in the lowest trace of Fig. 7. The individual bits are demarcated within each "pulse," and it is clear both that there tend to be significant run-lengths of bits within each pulse (due to the lowpass aspect), but also that these pulses themselves have an oscillatory character (due to the highpass aspect). These two tendencies suggest modelling IrisCode bit streams as a "sticky oscillator" Markov process.

III. Markov Generative Model of IrisCodes

The distinctive characteristics of IrisCode bit streams as seen in Fig. 7 can be effectively captured in a Hidden Markov Model (HMM) having two states and a single parameter, as depicted in Fig. 8. It may emit a '0' in State 0 or a '1' in State 1, each with probability α, and return to the same state; or with probability $1-\alpha$ it emits the other symbol and switches to the other state. Thus it can be "sticky" or "bouncy," two forms of predictability, depending on α. For $0.5 < \alpha < 1$ it is "sticky," with increasing run-lengths of the same bit as α gets larger. For $0 < \alpha < 0.5$ this two-state Markov process resists same-bit runs, alternating more regularly as α gets smaller. Regardless of α, we cannot derive the states from the outputs because either bit can be emitted in either state, and hence this Markov process is an HMM. In the case that $\alpha = 0.5$ it reduces to a simple one-state uncorrelated Bernoulli process, having a geometric distribution of run-lengths. For all values of the parameter $0 < \alpha < 1$, both bits are emitted with equal probability overall.

In order to test this model against real-world data, we begin with 11.5 million actual comparisons between IrisCodes generated from non-mated eyes, whose distribution of HD scores is shown in the upper panel of Fig. 9. Such actual IrisCode "imposter" distributions have been presented many times in the literature, and they are always very well-fitted by a fractional binomial distribution like the solid curve in both panels for IrisCode comparisons in a single orientation, or by the derived extreme-value variant of this distribution if only the best match (lowest score) after comparisons in several relative orientations is kept. The probability density function plotted as the solid curve fitting the empirical scores from comparisons

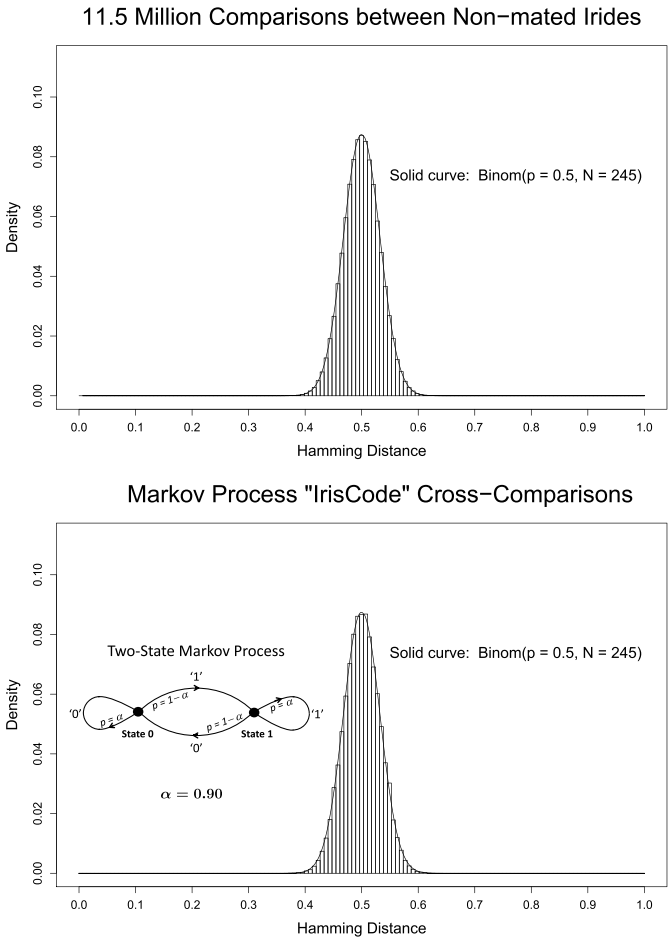

Fig. 9. Upper panel: distribution of actual HD scores obtained when IrisCodes from different eyes are compared, without multiple rotations. Solid curve is (3). Lower panel: Distribution of HD scores obtained between synthetic IrisCodes generated by the HMM shown in Fig. 8, with parameter $\alpha = 0.90$, without multiple rotations. Solid curve is the same (3) as in the upper panel.

between different eyes without multiple rotations is:

$$f_0(x) = \frac{N!}{m!(N-m)!}\ p^m(1-p)^{(N-m)} \tag{3}$$

where x here is the fractional HD score, the fraction of bits that disagreed between two IrisCodes from different eyes. More generally such a binomial distribution describes the fraction $x = m/N$ of Bernoulli trials in which one outcome (say "heads" in coin tosses) occurred, namely m out of N trials, and p is the probability of any single such outcome. For this empirical data, the parameters that produced the fitting binomial were $p = 0.5$ and $N = 245$.

The lower panel of Fig. 9 shows the distribution of HD scores obtained by cross-comparisons among 1,000 synthetic IrisCodes generated by the HMM depicted in Fig. 8. Each IrisCode had 1,536 unmasked bits after 25% of the 2,048 bits computed were deemed "unreliable", emulating the standard process of bit selection when encoding real iris data. Because the positions of those 25% discarded bits are randomly distributed, only 1,153 bits were mutually unmasked usually in any given pair of IrisCodes being compared. In the generative Markov Process producing these IrisCodes, various values of

the transition probability parameter α were tried until arriving at the value $\alpha = 0.9$ which perfectly matches the distribution obtained using actual iris images. In Fig. 9, the binomial density function that fits both score distributions (upper panel for actual eyes, lower panel for HMM-generated IrisCodes) is exactly the same curve. Therefore, this paper asserts that researchers could now generate large databases of synthetic IrisCodes having appropriate statistics by running the HMM of Fig. 8 with $\alpha = 0.9$ for 2,048 bit emissions per IrisCode, for purposes such as research on matching engines, indexing schemes, and the probabilities of extreme encounters.

IV. Why IrisCodes Have Proven so Resistant to False Matches

In order to test the model against vastly larger datasets of scores, such as the 1.2 trillion iris comparisons performed by the US National Institute of Standards and Technology (NIST), we must incorporate the multiple rotations comparisons which increase False Match probability (FMR). Because it cannot be known precisely what was the amount of head tilt, camera tilt, or eye rotation (cyclovergence) when the IrisCodes were obtained, it is necessary to make comparisons over a reasonable range of k relative tilts (rotations) between every pair of IrisCodes, keeping the best match as their similarity score. This generates an *extreme value distribution* that is skewed towards lower HD scores, even for unrelated eyes, because of the increased opportunities to get a closer match just by chance. In effect the search space is k times larger, and the net False Match probability (for a given threshold) is therefore almost k times larger. Most current public deployments of iris recognition use $k = 7$ relative tilt angles, but a few go as far as $k = 21$ relative tilt angles when handheld cameras are used.

Let $f_0(x)$ be the density distribution obtained for match scores x between different irides after comparing them in only a single relative orientation. For example, $f_0(x)$ might be the fractional binomial defined in (3) and plotted in Fig. 9, or any other probability distribution on $x \geq 0$. Then $F_0(x)$, the cumulative of $f_0(x)$ from 0 to x, becomes the probability of getting a False Match in such a test when using the criterion that any dissimilarity score $\leq x$ is accepted as a match:

$$F_0(x) = \int_0^x f_0(x)dx \tag{4}$$

or, equivalently,

$$f_0(x) = \frac{d}{dx}F_0(x) \tag{5}$$

Clearly, then, the probability of *not* making a False Match when using decision criterion x is $1 - F_0(x)$ after a single test, and it is $[1 - F_0(x)]^k$ after carrying out k such tests independently at k different relative orientations. It follows that the probability of a False Match after a "best of k" test of agreement, when using criterion x, regardless of the actual form of the raw unrotated distribution $f_0(x)$, is:

$$F_k(x) = 1 - [1 - F_0(x)]^k \tag{6}$$

TABLE I

False Match Rates Predicted by Eqt (6) and as Measured by NIST [15] With 1.16 Billion Iris Comparisons, and [16] With 1.2 Trillion Iris Comparisons

HD Criterion	*Predicted FMR,* (6)	*NIST*[15],[16] *Measured FMR*
0.36	1 in 24,000	1 in 25,000
0.35	1 in 110,000	1 in 71,000
0.34	1 in 556,000	1 in 476,000
0.33	1 in 3 million	1 in 3.4 million
0.32	1 in 20 million	1 in 24 million
0.31	1 in 137 million	1 in 165 million
0.30	1 in 1.1 billion	1 in 2 billion
0.29	1 in 9 billion	(not measured)
0.28	1 in 92 billion	1 in 40 billion

and the expected density $f_k(x)$ associated with this cumulative is:

$$\begin{aligned} f_k(x) &= \frac{d}{dx}F_k(x) \\ &= kf_0(x)\,[1 - F_0(x)]^{k-1} \end{aligned} \tag{7}$$

Detailed numerical tabulations of these probability densities $f_0(x)$ and $f_k(x)$, their associated cumulatives $F_0(x)$ and $F_k(x)$, and the $\log_{10}$ of False Match probability as a function of HD acceptance criterion x, for $f_0(x)$ defined as in (3), are available online at [14] in both human-readable and machine-readable formats.

Table I shows how remarkably well the predictions of this model compare with actual accuracy results reported in independent tests of the IrisCode by NIST, over a range of more than six log units of FMR variation determined by the decision criterion. IREX-I [15] performed 1.16 billion, and IREX-III [16] performed 1.2 trillion, IrisCode comparisons. To obtain such large numbers of pairings using *intra*-dataset comparisons, as was done in [6], is risky because of the inevitability of biographical ground-truth errors which can dominate estimates of accuracy. Databases acquired within universities using student populations recruited by payment naïvely provide an incentive for such ground-truth errors. As the director of one famous such effort eventually conceded, after first reporting many False Matches later shown to be illusory: "Clearly we were getting scammed by some of our student volunteers; (being paid to enroll) they were changing names and coming through multiple times." Similar incentives for persons to enroll under multiple identities exist with detainee and expellee populations. The consequence of even a single such subject having two identities when N subjects are enrolled for full intra-dataset comparisons, is that the estimated FMR can never be better than $2/N^2$. The measured threshold calibration of FMR such as shown in Table I would approach a floor FMR that cannot be reduced by any reasonable change in threshold, and indeed NIST [15] demonstrated this problem for intra-dataset comparisons.

Instead, performing *inter*-dataset comparisons can avoid the contaminating effect of biographical ground-truth errors. If two disjoint populations, of sizes say N and M in geographically remote regions can be biometrically enrolled, then $N \times M$ inter-comparisons become possible without illusory

False Matches. NIST [16] acquired enrollment datasets of two populations "very well separated geographically and occupationally," one having 3.9 million iris images as the gallery and the other having 315,000 iris images as probes used to search against this entire gallery, asserting that "the likelihood of co-membership is considered to be identically zero." Thereby NIST [16] was able to perform $N \times M = 1,228$ billion, or 1.2 trillion, IrisCode comparisons leading to the results shown in Table I for various HD threshold criteria. The close confirmation of theory, over more than six log units, is striking. At the bottom of the table, at criterion HD = 0.28 meaning that 28% of the bits in two IrisCodes are allowed to disagree while still accepting them as a match, even the factor of two between the prediction of FMR = 1 in 92 billion and the NIST observation of FMR = 1 in 40 billion ([16, p. 61]) remains impressively concordant since these two FMR probabilities are respectively the -10.96^{th} and -10.60^{th} powers of ten. A further important feature of Table I is that for HD criteria below about 0.33, each percentile point reduction in HD (*e.g.* from 0.31 to 0.30) brings almost another order of magnitude reduction in FMR. This high-leverage consequence of IrisCode bit combinatorics is what enables massive populations like India's to be enrolled, with quadratic cross-comparisons for de-duplication, without drowning in False Matches.

The False Match error rates presented in Table I, with the model closely confirmed by the recent NIST large-scale tests [15], [16], are in fact very close to predictions made by this author [5] in 2003. The theoretical [5, Table 1, p. 287] very closely parallels Table I in this paper. But those performance predictions were treated dismissively and incredulously by the biometrics community because such FMR performance was unheard of in other biometrics. Some researchers [17] at NIST published papers that were even contemptuous of these claims, saying instead that iris recognition was no more powerful than face recognition. They compared both technologies at the extremely undemanding criterion of FMR = 0.001, not realising that the DET (Decision Error Tradeoff) curves for iris are so flat that the FMR can be reduced by many orders of magnitude through small reductions in threshold (see Table I) while having only minuscule impact on FnMR. The slope of such error trade-off curves is called the *likelihood ratio,* and it equals the ratio of the two probability density distributions (for same-person and different-person comparisons) at the chosen decision criterion. As confirmed in [16], the IrisCode DET slope is so small that its FMR can be lowered by a factor of 10,000 to 100,000 while not even doubling the FnMR.

Comparing face and iris in identical one-to-many protocols, later researchers at NIST [16] recently concluded: "The shape of the respective DETs indicates that iris will give at least 100,000 times fewer false positives than face, for an equal false negative identification rate." Likewise, at a fixed FMR, "iris gives a factor of ten fewer misses than face. Two-iris operation would double this improvement." ([16, p. 8]). It is very gratifying that these recent confirmations of the 2003 predictions [5] for IrisCode accuracy have in fact come from NIST itself. They confirm the key biometric role of entropy, as the source of biometric discriminating power. Later we will study this further after using white noise analysis to separate the properties of the "signal" from those of the "channel."

V. The Biometric "Birthday Problem"

The familiar "birthday problem" asks how many persons chosen at random are needed before it becomes more likely than not that at least one pair in this group share a birthday. As N persons have $N(N-1)/2$ possible pairings, and any given pair has probability $365/366$ of *not* sharing birthdays, the problem requires only that $(365/366)^{N(N-1)/2} < 0.5$ for birthday collisions to be more likely than not. Some people find it counterintuitive and surprising that this condition is passed when there are as few as $N = 23$ persons in the group.

An analogous problem exists for biometrics: given a False Match probability FMR (determined by a chosen threshold and how discriminating the modality is), how large can a population become before it is likelier than not that at least two persons in the group collide biometrically? For face recognition with its benchmark FMR = 0.001 test standard, the answer is again easy to calculate and perhaps surprising: collisions are to be expected once there are just $N = 38$ persons. The general solution is easy to derive for any FMR that is small: biometric collisions are likelier than not among a group of N persons once their number is $N > \sqrt{1.386/\text{FMR}}$. This calculation should be done for every biometric modality at whatever FMR is considered reasonable and achievable for that modality. It is the reason why the "astronomical" odds against False Matches as tabulated in [14] as per (6) and in Table I as confirmed empirically by NIST [16], are important in real deployments and, when amplified quadratically by using both eyes, are critical for de-duplication cross-comparison checks in national ID projects such as UIDAI [2].

VI. White Noise Analysis of the IrisCode

We turn now to another important method of information theory: white noise analysis. The IrisCode can be regarded as a kind of a channel, and the capacity of a channel is defined by the maximum of the mutual information between input and output over all possible input distributions. It is known [1], [18] that the distribution which has maximum entropy for any given variance is Gaussian white noise: a random signal whose samples are independent (hence uncorrelated) and identically distributed with values having a Gaussian probability distribution, and whose Fourier power spectral density is uniform (hence "white" noise). We wish to study how the IrisCode encodes and matches such signals. One purpose for this is to understand what portion of the correlations observed within an IrisCode computed from a real iris (see Fig.s 4 and 7) arise from iris texture itself, given the limited entropy of natural anatomical structures, and what portion is imposed by the IrisCode as a channel even with uncorrelated random input. Therefore, 500 artificial iris images such as seen in Fig. 10 were synthesised using pixel values sampled from a Gaussian white noise process ($\mu = 128, \sigma = 20$). The histograms in the lower right show the count of pixels within an artificial iris at each grey level from 0 to 255. All 500 such iris images were processed in the normal way to compute

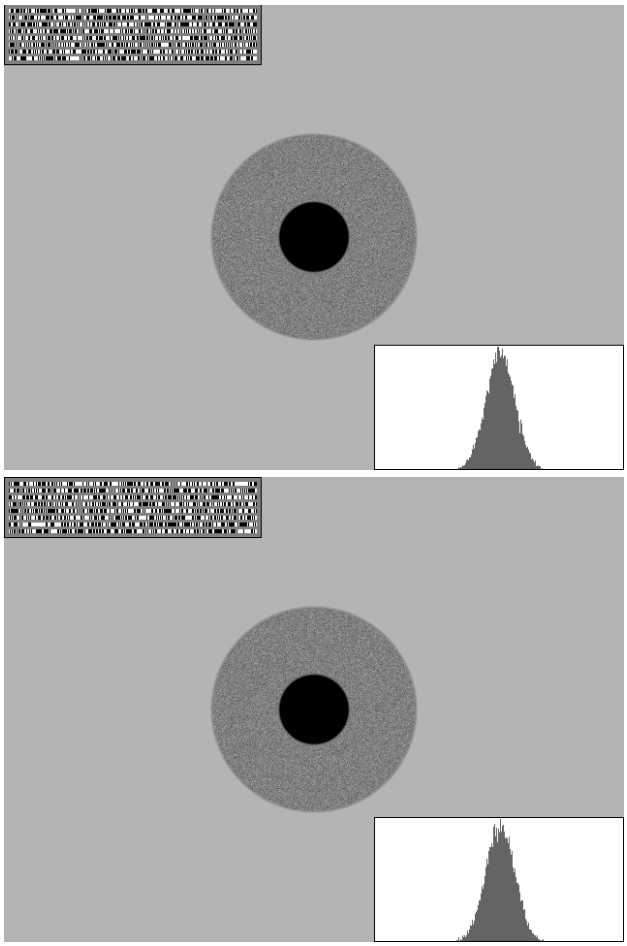

Fig. 10. Two examples of the 500 white noise iris images synthesized to analyse the capacity of the IrisCode as a channel, without entropy reductions from anatomical correlations at input. Iris pixel distributions are Gaussian.

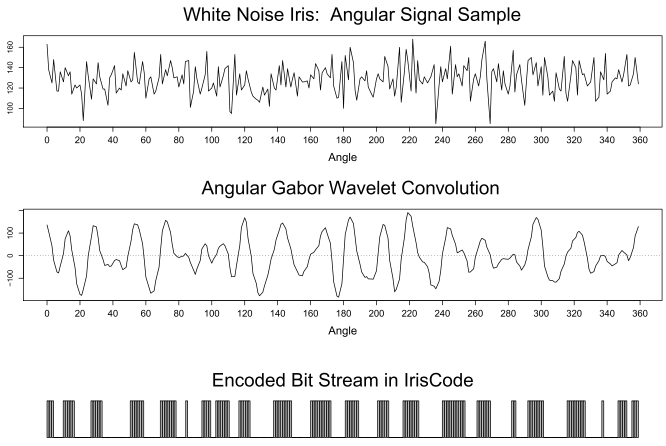

Fig. 11. Pixel values extracted from one synthetic white noise iris along a mid-radius circumference (upper trace). Output from Gabor wavelet encoder (same as for Fig. 7, middle trace) extracts just a middle-low frequency band.

their IrisCodes as shown in the upper left, and then each of these was matched against all of the others, leading to an "imposters" distribution of $500 \times 499/2 = 124,750$ HD scores.

The upper trace in Fig. 11 plots the pixel values sampled from one white noise iris along a single (mid-radius) circumference, just as was done in Fig. 7 for a real iris. It obviously contains a higher range of frequencies than the natural iris, but it lacks the low frequency gradients created by illumination geometry. The middle trace showing output from the bandpass Gabor wavelet encoder is of course quite similar to that for the natural iris, as the same frequency band has been extracted, but both this trace and the quantised output bit stream in the lower trace suggest that slightly higher frequency content remains encoded. As before, the individual bits within each "pulse" are demarcated, and they resemble a "sticky oscillator."

It is important to note that the choice of variance for the Gaussian noise process ($\sigma = 20$) is immaterial for the analysis and results presented here, because σ merely determines the amplitude, or the contrast, of the noise. It controls the widths of the pixel histograms embedded within Fig. 10, and it scales the amplitude of the two signal traces in Fig. 11. But no changes in σ can change the locations of the zero-crossings in the second trace, which determine the output bit sequence. It is also important to recall that the defining integral (2) which sets bits in an IrisCode is a two-dimensional polar integral, so it is summing convolutional inner products in both the radial and angular coordinates. Thus the structure of the HMM bit stream is inherently modelling radial as well as angular correlations within regions of an iris.

The actual distribution of HD scores obtained when all 500 white noise iris images were compared against each other (without relative rotations to seek best matches) is plotted in the upper histogram of Fig. 12. The solid curve that fits the distribution perfectly is again the fractional binomial of (3) but in this case with parameters $p = 0.5$ and $N = 337$, and thus the distribution is noticeably narrower than for natural IrisCodes presented in Fig. 9 which used $N = 245$. These values for N were chosen by measuring the standard deviation σ for these distributions, observing their mean is $p = 0.5$ and then noting that for a fractional binomial distribution, $\sigma = \sqrt{p(1-p)/N}$. The lower panel of Fig. 12 shows that exactly the same distribution (fit by the same solid curve) is obtained when comparing artificial IrisCodes created using bit sequences emitted by the HMM model of Fig. 8, masking a random quartile of bits deemed "unreliable" in each IrisCode (so that typically 1,153 bits were mutually unmasked in an IrisCode pair being compared), but not performing rotations to seek best matches. The standard score normalisation [6, eq. (6)] that compensates for the number of bits mutually available for comparison between any two IrisCodes is applied. Most importantly, the value of the parameter α needed in the two-state Markov Process to produce this distribution, as fit by the solid curve, is $\alpha = 0.867$ which makes it a rather less "sticky" oscillator than the HMM needed to emulate the IrisCodes actually produced by natural iris images (Fig. 9). Now we are finally in a position to estimate the entropy of these processes and to estimate how much of the intrinsic capacity of the IrisCode is used when encoding natural iris texture.

VII. Conclusion: Capacity of the IrisCode

If we accept that the HMM of Fig. 8 appears well able to emulate IrisCodes whether they are computed from real iris

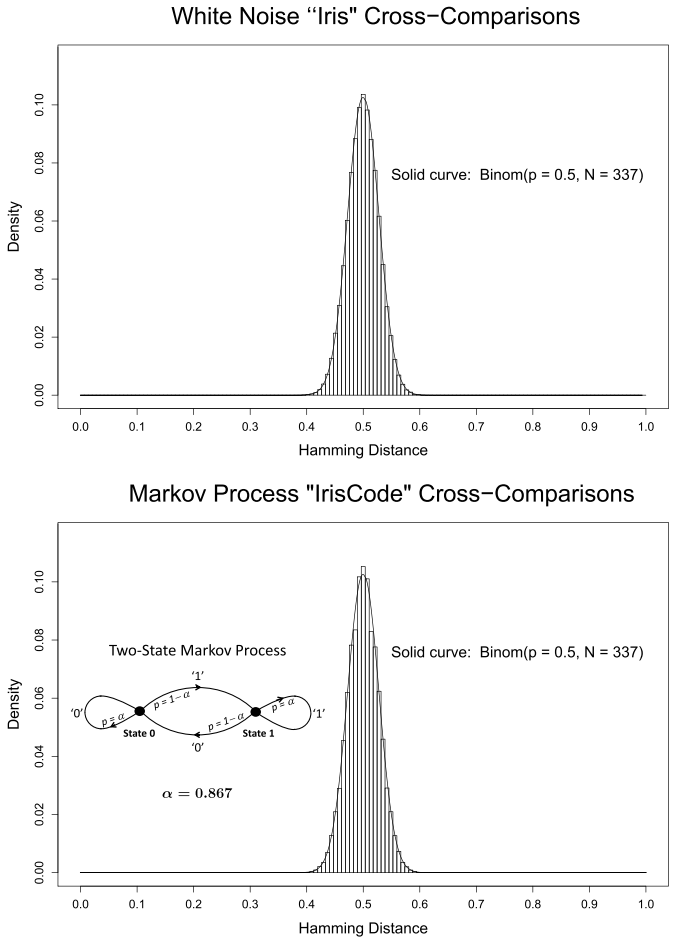

Fig. 12. Upper panel: distribution of HD scores obtained when the IrisCodes computed for 500 white noise iris images are cross-compared, without multiple rotations. Solid curve is (3). Lower panel: Distribution of HD scores obtained between synthetic IrisCodes generated by the HMM of Fig. 8, with parameter $\alpha = 0.867$, without multiple rotations. Solid curve is the same (3) as in the upper panel.

images (Fig. 9) or from white noise iris images (Fig. 12), then it is interesting to calculate the entropy of this generative process, in bits of entropy per bit emitted, for both cases.

The overall entropy of a multi-state Markov Process is calculated by combining the entropies of all the states, weighted by their respective state occupancy probabilities [18]. Because both switches in the HMM of Fig. 8 are equiprobable and the flow paths between these states are symmetrical, both states clearly have equiprobable occupancy $p(\text{State } 0) = p(\text{State } 1) = 0.5$ regardless of the value of α in $0 < \alpha < 1$. From the definition (1) of entropy in terms of event probabilities for $p_0 = \alpha$, the entropy of State 0 is $H(\alpha) = -\alpha \log_2(\alpha) - (1-\alpha)\log_2(1-\alpha)$. State 1 also has the same entropy. Thus the overall entropy $H(\alpha)$ of this Markov source (in bits per bit emitted) is yet again this same expression:

$$H(\alpha) = -\alpha \log_2(\alpha) - (1-\alpha)\log_2(1-\alpha) \tag{8}$$

For a Bernoulli process (like coin-tossing) whose two outcomes have probabilities p and $1-p$, Fig. 13 is a plot of the Shannon entropy (1) as a function of p. In (8) we saw that the entropy $H(\alpha)$ of the "sticky oscillator" HMM which

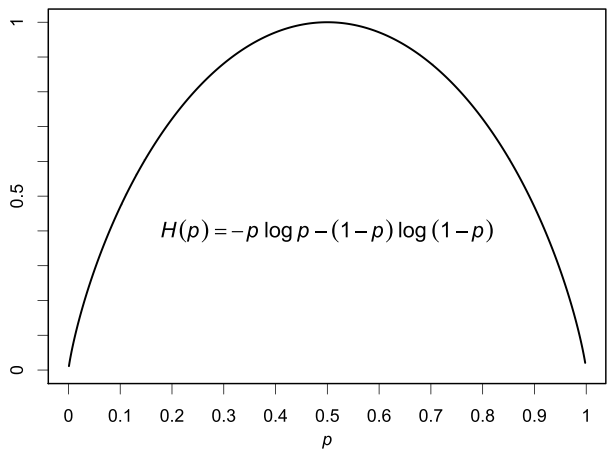

Fig. 13. Plot of (1) for the entropy of a Bernoulli process as a function of its outcome probability p. This is also a plot of (8) for the entropy of the two-state HMM in Fig. 8 whose transition probability parameter is $p = \alpha$.

models IrisCodes very well also takes this same functional form, with $p = \alpha$. Note that if $\alpha = 0.5$ then the HMM of Fig. 8 simplifies to a single-state uncorrelated Bernoulli process having maximum possible entropy: it produces 1 bit of entropy per bit emitted. For values of α moving away from 0.5 in either direction, the entropy of the process declines, approaching $H(\alpha) = 0$ as it tends either towards complete "stickiness" ($\alpha \to 1$) or complete "bounciness" ($\alpha \to 0$), both of which are forms of predictability.

We conclude from Fig. 9 and (8) that IrisCodes computed from real iris images contain an entropy of $H(0.9) = 0.469$ bits per bit. But the capacity of the IrisCode as revealed in Fig. 12 when encoding white noise iris images such as shown in Fig. 10, which have the maximum possible entropy for any given variance, is $H(0.867) = 0.566$ bits per bit. The difference between 0.566 and 0.469 bits of entropy per encoded bit is a reflection of the existence of correlations (non-randomness) within the anatomy of a natural iris. It is the IrisCode's "unused capacity" for discrimination when encoding natural iris images.

A quantitative way to understand why the large entropy of the IrisCode is the origin of its extreme resistance to False Matches is to consider the monstrously large factorial terms that dominate the combinatorial part of (3). There is a good analogy between XOR-ing the bits in IrisCodes from different eyes and tossing a fair coin many times in a run (say N tosses), because all possible sequences are equiprobable (namely every sequence has probability 2^{-N}) but there are vastly more outcome sequences containing relatively balanced mixes of heads and tails than less balanced mixes. This distribution of fractions is exactly the fractional HD distribution fitted by the solid curves in Fig.s 9 and 12, one requiring $N = 245$ and the other $N = 337$, and their tails attenuate extremely rapidly. For example, if tossing a fair ($p = 0.5$) coin $N = 245$ times in a row, the tabulation of $F_0(x)$ in [14] as per (4) reveals that getting fewer than 27% heads is ten times less likely than getting fewer than 28% heads. The extremely rapid attenuation of the tail of this combinatorial distribution, namely by a factor of ten for a mere one percentile point change in the

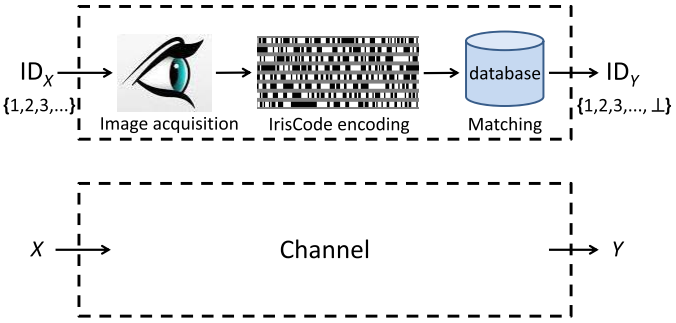

Fig. 14. We can regard the entire process of biometric identification as constituting a kind of noisy channel, with a presenting actual identity X as the input, and an inferred identity Y as the output from biometric matching (including $\perp$ for "unknown" or "null identity").

balance of outcomes (from 28% heads to 27% heads), mirrors the behaviour seen in Table I as confirmed by NIST, and it summarises the importance of large entropy for biometric discriminating power.

In future work we will regard the human presentation and the image acquisition process as also constituting parts of the channel, so the input as suggested in Fig. 14 is an identity rather than an image, and the output is also an identity (possibly including $\perp$ for "unknown"). Such an approach depends on the "authentics" distribution and on how image quality affects the probability of False non-Matches, whereas the present work focused on the discriminating capacity of the IrisCode: the properties that determine its famous resistance to False Matches. The entire process of iris recognition from presentation and image acquisition to matching and decision-making is a channel between input identities and output identities, whose mutual information we seek to maximise.

ACKNOWLEDGMENTS

The author would like to thank Markus Kuhn for generating the white noise images, and Cathryn Downing for running the Markov process simulations.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.
[2] *Unique Identification Authority of India: Dashboard Showing Enrollment Progress, Updated Weekly*. [Online]. Available: https://portal.uidai.gov.in/uidwebportal/dashboard.do, accessed Apr. 27, 2015.
[3] S. Nadhamuni, "The unique identification authority of India," in *Proc. Int. Conf. Biometrics*, New Delhi, India, Mar./Apr. 2012.
[4] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 11, pp. 1148–1161, Nov. 1993.
[5] J. Daugman, "The importance of being random: Statistical principles of iris recognition," *Pattern Recognit.*, vol. 36, no. 2, pp. 279–291, 2003.
[6] J. Daugman, "Probing the uniqueness and randomness of IrisCodes: Results from 200 billion iris pair comparisons," *Proc. IEEE*, vol. 94, no. 11, pp. 1927–1935, Nov. 2006.
[7] A. W. K. Kong, D. Zhang, and M. S. Kamel, "An Analysis of IrisCode," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 522–532, Feb. 2010.
[8] A. W. K. Kong, "A statistical analysis of IrisCode and its security implications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 513–528, Mar. 2015.
[9] J. Daugman, "Statistical richness of visual phase information: Update on recognizing persons by iris patterns," *Int. J. Comput. Vis.*, vol. 45, no. 1, pp. 25–38, Oct. 2001.
[10] K. P. Hollingsworth, K. W. Bowyer, and P. J. Flynn, "Improved iris recognition through fusion of Hamming distance and fragile bit distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2465–2476, Dec. 2011.
[11] H. Proença, "Iris recognition: What is beyond bit fragility?" *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 2, pp. 321–332, Feb. 2015.
[12] J. R. Matey, R. Broussard, and L. Kennell, "Iris image segmentation and sub-optimal images," *Image Vis. Comput.*, vol. 28, no. 2, pp. 215–222, 2010.
[13] C. Rathgeb, A. Uhl, and P. Wild, "Shifting score fusion: On exploiting shifting variation in iris recognition," in *Proc. ACM Symp. Appl. Comput.*, New York, NY, USA, 2011, pp. 3–7.
[14] *Table of Probability Densities and Their Cumulatives*. [Online]. Available: http://www.CL.cam.ac.uk/users/jgd1000/IrisCumulatives.pdf and http://www.CL.cam.ac.uk/users/jgd1000/binom245data.machine.readable, accessed Apr. 30, 2015.
[15] P. Grother, E. Tabassi, G. W. Quinn, and W. Salamon, "IREX I: Performance of iris recognition algorithms on standard images," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. 7629, Oct. 2009.
[16] P. J. Grother *et al.*, "IREX III: Performance of iris identification algorithms," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. 7836, Apr. 2012.
[17] E. M. Newton and P. J. Phillips, "Meta-analysis of third-party evaluations of iris recognition," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 39, no. 1, pp. 4–11, Jan. 2009.
[18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.

John Daugman received his degrees from Harvard University and then taught at Harvard before coming to the University of Cambridge, where he is a Professor of Computer Vision and Pattern Recognition. He has held the Johann Bernoulli Chair of Mathematics and Informatics at the University of Groningen, and the Toshiba Endowed Chair at the Tokyo Institute of Technology. His areas of research and teaching at Cambridge include computer vision, information theory, and statistical pattern recognition. He is the inventor of iris recognition, and his algorithms are the core of all public operational deployments of the technology. Awards for his work in science and technology include the Information Technology Award, the Medal of the British Computer Society, the "Time 100" Innovators Award, and the Order of the British Empire. He has been elected as a fellow of the Royal Academy of Engineering, the Institute of Mathematics and its Applications, and the British Computer Society. In 2013, he was inducted into the U.S. National Inventors Hall of Fame.