

New Methods in Iris Recognition

John Daugman

Abstract—This paper presents the following four advances in iris recognition: 1) more disciplined methods for detecting and faithfully modeling the iris inner and outer boundaries with active contours, leading to more flexible embedded coordinate systems; 2) Fourier-based methods for solving problems in iris trigonometry and projective geometry, allowing off-axis gaze to be handled by detecting it and “rotating” the eye into orthographic perspective; 3) statistical inference methods for detecting and excluding eyelashes; and 4) exploration of score normalizations, depending on the amount of iris data that is available in images and the required scale of database search. Statistical results are presented based on 200 billion iris cross-comparisons that were generated from 632 500 irises in the United Arab Emirates database to analyze the normalization issues raised in different regions of receiver operating characteristic curves.

Index Terms—Active contours, biometrics, Gabor wavelets, gaze correction, iris recognition, score normalization, texture.

I. INTRODUCTION

THE ANTICIPATED large-scale applications of biometric technologies such as iris recognition are driving innovations at all levels, ranging from sensors to user interfaces, to algorithms and decision theory. At the same time as these good innovations, possibly even outpacing them, the demands on the technology are getting greater. Today, many countries are considering or have even announced procurement of biometrically enabled national identity (ID) card schemes, one of whose purposes will be to detect and prevent multiple IDs. Achieving that purpose will require, at least at the time when cards are issued and IDs are registered, an offline “each-against-all” cross-comparison. In effect then the number of biometric comparisons that must be performed scales as the square of the population. The decision confidence levels that will be required to keep the false match rate (FMR) negligible, despite such vast numbers of opportunities to make false matches, can only be described as astronomical.

At the other end of the conceptual receiver operating characteristic (ROC) curve from the $FMR = 0$ limit, much greater demands are also being placed on the false non-match rate (FnMR). This is partly because national-scale deployments must be as inclusive as possible; hence, it is unacceptable to exclude members of outlier populations who, for various reasons, may have a nonstandard eye appearance (e.g., non-round iris, coloboma, oddly shaped pupil, drooping eyelids, or much eyelash occlusion) or who simply have difficulty

presenting to the camera (e.g., nystagmus or deviated gaze). The demands against false non-matches are also being raised by the development of more tolerant and fluid user interfaces, which aim to replace the “stop-and-stare” camera interface with iris recognition on the move, off-axis, and at a distance [9].

These two trends seem to require, paradoxically, that decision criteria be used which are simultaneously much more conservative and liberal than those that are presently deployed. The purpose of this paper is to present four new advances in iris recognition that aim to simultaneously improve at both extremes. Sections II–IV present new methods of image processing for iris segmentation, allowing flexible shapes and coordinate systems. The images used in these sections come from the National Institute of Standards and Technology (NIST) “Iris Challenge Evaluation” (ICE-1), and their file numbers are cited to facilitate comparative work. Sections V and VI analyze alternative score normalization rules that are adapted for differently sized databases, using results from 200 billion iris cross-comparisons.

II. ACTIVE CONTOURS AND GENERALIZED COORDINATES

Iris recognition begins with finding an iris in an image, demarcating its inner and outer boundaries at the pupil and sclera, detecting the upper and lower eyelid boundaries if they occlude, and detecting and excluding any superimposed eyelashes or reflections from the cornea or eyeglasses. These processes may collectively be called segmentation. Precision in assigning the true inner and outer iris boundaries, even if they are partly invisible, is important because the mapping of the iris in a dimensionless (i.e., size invariant and pupil dilation invariant) coordinate system is critically dependent on this. Inaccuracy in the detection, modeling, and representation of these boundaries can cause different mappings of the iris pattern in its extracted description, and such differences could cause failures to match.

It is natural to start by thinking of the iris as an annulus. Soon, one discovers that the inner and outer boundaries are usually not concentric. A simple solution is then to create a nonconcentric pseudo-polar coordinate system for mapping the iris, relaxing the assumption that the iris and pupil share a common center and requiring only that the pupil is fully contained within the iris. This “doubly dimensionless pseudo-polar coordinate system” was the basis of my original paper on iris recognition [2] and patent [3], and this iris coordinate system was incorporated into International Organization for Standardization (ISO) Standard 19794-6 for iris data [7]. Soon one also discovers that, often, the pupil boundary is noncircular, and usually, the iris outer boundary is noncircular. Performance in iris recognition is significantly improved by relaxing both of those assumptions, replacing them with more disciplined

Manuscript received September 20, 2006. This paper was recommended by Guest Editor K. W. Bowyer.

The author is with the Computer Laboratory, University of Cambridge, CB3 0FD Cambridge, U.K.

Digital Object Identifier 10.1109/TSMCB.2007.903540

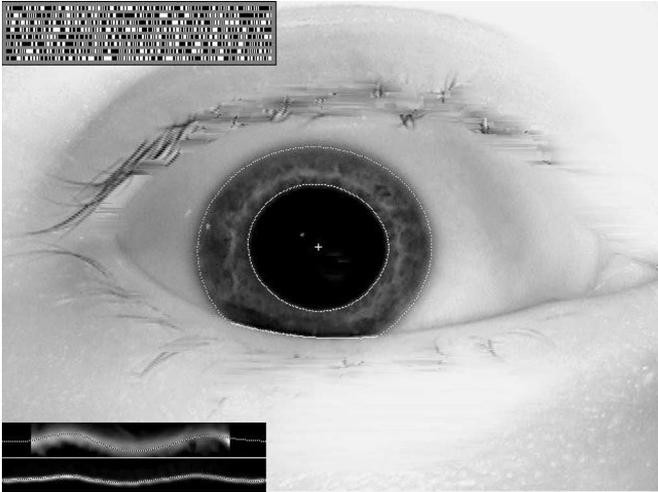


Fig. 1. Active contours enhance iris segmentation, because they allow for noncircular boundaries and enable flexible coordinate systems. The box in the lower left-hand corner shows curvature maps for the inner and outer iris boundaries, which would be flat and straight if they were circles. Here, the outer boundary (upper plot) is particularly noncircular. Dotted curves in the box and on the iris are Fourier series approximations. This iris is ICE-1 file 239261.

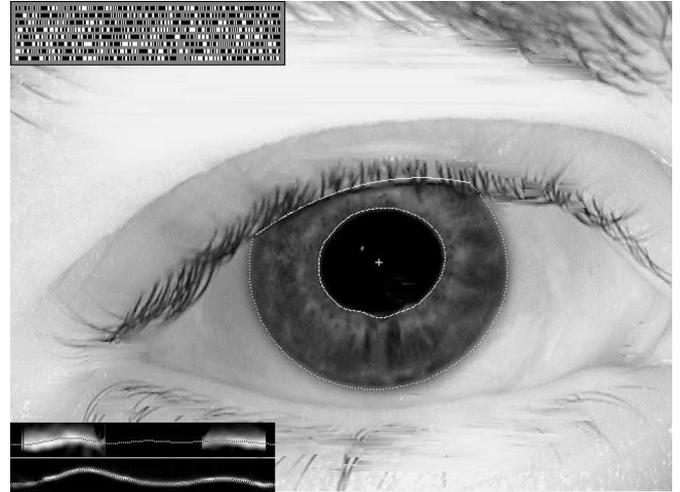


Fig. 2. Active contours enhance iris segmentation, because they allow for noncircular boundaries and enable flexible coordinate systems. The box in the lower left-hand corner shows curvature maps for the inner and outer iris boundaries, which would be flat and straight if boundaries were circles. Here, the pupil boundary (lower plot) is particularly noncircular. Dotted curves in the box and on the iris are Fourier series approximations. This iris is ICE-1 file 240461.

methods for faithfully detecting and modeling those boundaries whatever their shapes are and defining a more flexible and generalized coordinate system on their basis.

Because the iris outer boundary is often partly occluded by eyelids, and the iris inner boundary may be partly occluded by reflections from illumination, and sometimes both boundaries also by reflections from eyeglasses, it is necessary to fit flexible contours that can tolerate interruptions and continue their trajectory across them on a principled basis, which is somehow driven by the data that exist elsewhere. A further constraint is that both the inner and outer boundary models must form closed curves. A final goal is that we would like to impose a constraint on smoothness based on the credibility of any evidence for nonsmooth curvature.

An excellent way to achieve all of these goals is to describe the iris inner and outer boundaries in terms of “active contours” based on discrete Fourier series expansions of the contour data. By employing Fourier components whose frequencies are integer multiples of $1/(2\pi)$, closure, orthogonality, and completeness are ensured. Selecting the number of frequency components allows control over the degree of smoothness that is imposed and over the fidelity of the approximation. In essence, truncating the discrete Fourier series after a certain number of terms amounts to low-pass filtering the boundary curvature data in the active-contour model.

These methods are illustrated in Figs. 1 and 2. The lower left-hand corner of each figure shows two “snakes,” each consisting of a fuzzy ribbon-like data distribution and a dotted curve, which is a discrete Fourier series approximation to the data, including continuation across gap interruptions. The lower snake in each snake box is the curvature map for the pupil boundary, and the upper snake is the curvature map for the iris outer boundary, with the endpoints joining up at the six o’clock position. The interruptions correspond to detected occlusions by eyelids, which are indicated by separate splines in both images, or by specular reflections. The data plotted as the gray

level for each snake is the image gradient in the radial direction. Thus, the relative thickness of each snake roughly represents the sharpness of the corresponding radial edge. If an iris boundary were well-described as a circular edge, then the corresponding snake in its box should be flat and straight. In general, this is not the case.

The dotted curve that is plotted within each snake, as well as superimposed on the corresponding loci of points in the iris image, is a discrete Fourier series approximation to the data. (In both figures, detected eyelid occlusions are also demarcated by white pixels, and they interrupt the corresponding outer boundary data snake, although the estimated contour continues through such interruptions.) The estimation procedure is to compute a Fourier expansion of the N regularly spaced angular samples of radial gradient edge data $\{r_\theta\}$ for $\theta = 0$ to $\theta = N - 1$. A set of M discrete Fourier coefficients $\{C_k\}$, for $k = 0$ to $k = M - 1$, is computed from the data sequence $\{r_\theta\}$ as follows:

$$C_k = \sum_{\theta=0}^{N-1} r_\theta e^{-2\pi i k \theta / N}. \quad (1)$$

Note that the zeroth-order coefficient or “DC term” C_0 extracts information about the average curvature of the (pupil or outer iris) boundary, or in other words, about its radius when it is approximated as just a simple circle.

From these M discrete Fourier coefficients, an approximation to the corresponding iris boundary (now without interruptions and at a resolution determined by M) is obtained as the new sequence $\{R_\theta\}$ for $\theta = 0$ to $\theta = N - 1$, which is expressed as follows:

$$R_\theta = \frac{1}{N} \sum_{k=0}^{M-1} C_k e^{2\pi i k \theta / N}. \quad (2)$$

As is generally true of active-contour methods [1], [8], there is a tradeoff between how precisely one wants the model to fit all the data (improved by increasing M) versus how much one wishes to impose constraints such as keeping the model simple and of low-dimensional curvature (achieved by reducing M ; e.g., $M = 1$ enforces a circular model). Thus, the number M of activated Fourier coefficients is a specification for the degrees of freedom in the shape model. It has been found that a good choice of M for capturing the true pupil boundary with appropriate fidelity is $M = 17$, whereas a good choice for the iris outer boundary where the data is often much weaker is $M = 5$. It is also useful to impose monotonically decreasing weights on the computed Fourier coefficients $\{C_k\}$ as further control on the resolution of the approximation $\{R_\theta\} \approx \{r_\theta\}$, which amounts to low-pass filtering the curvature map in its Fourier representation. Altogether, these manipulations, particularly the two different choices for M , implement the computer vision principle that strong data (the pupil boundary) may be modeled with only weak constraints, whereas weak data (the outer boundary) should be modeled with strong constraints.

The active-contour models for the inner and outer iris boundaries support an isometric mapping of the iris tissue between them, regardless of the actual shapes of the contours. Suppose the contour model for the pupillary boundary consists of Cartesian coordinates $(x_p(\theta), y_p(\theta))$ with arc parameter $\theta \in [0, 2\pi]$ and the outer boundary of the iris at the sclera is described by the contour model $(x_s(\theta), y_s(\theta))$. Then, a shape-flexible, size-invariant, and pupil-dilation-invariant dimensionless coordinate system for the iris portion of the image $I(x, y)$ can be represented by the following normalized mapping:

$$I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta) \quad (3)$$

where the dimensionless parameter $r \in [0, 1]$ spans the unit interval, and

$$\begin{bmatrix} x(r, \theta) \\ y(r, \theta) \end{bmatrix} = \begin{bmatrix} x_p(\theta) & x_s(\theta) \\ y_p(\theta) & y_s(\theta) \end{bmatrix} \begin{bmatrix} 1 - r \\ r \end{bmatrix}. \quad (4)$$

The execution time for the entire subroutine that fits active contours to both the inner and outer iris boundaries is just 3.5 ms on a 3-GHz PC with optimized code. The benefit of the new adaptive coordinate system based on the active contours may be gauged by the improvement that it offers in recognition performance on difficult image databases. The NIST ICE-1 iris database contains many difficult images, producing a high false reject rate (FRR), which degrades the equal error rate (EER). Algorithms that yielded an EER of 1% (i.e., $\text{EER} = 0.01$) when using enforced circular models improved tenfold to an EER of 0.1% (i.e., $\text{EER} = 0.0011$) on the same database by adopting this active-contours approach instead.

Because of the flexibility of this method and its ability to establish a continuous deformation mapping between arbitrary shapes, we may call this a “generalized embedded coordinate system,” or perhaps, more evocatively, we may use the name “Faberge coordinates” [10].

III. FOURIER-BASED TRIGONOMETRY AND OFF-AXIS GAZE

A limitation of current iris recognition cameras is that they require an on-axis image of an eye, which is usually achieved through what may be called a “stop-and-stare” interface, in which a user must align her optical axis with the camera’s optical axis. This is not as flexible or fluid as it might be. Moreover, sometimes, the standard cameras acquire images for which the on-axis assumption is not true. For example, the NIST iris images that were made available and used for testing in ICE-1 contained several images with very deviated gaze, probably because the user’s gaze was distracted by an adjacent monitor.

The on-axis requirement can be relaxed by correcting the projective deformation of the iris when it is imaged off-axis, provided that one can reliably estimate the actual parameters of gaze. Dorairaj *et al.* [5] approached this by seeking the gaze parameters that optimize the value of an integro-differential operator [2], [4], which detects circular boundaries. The gaze parameters that we seek include two spherical angles for eye pose, but the projective geometry also depends on the distance between the eye and the camera, which may be unknown, and on the surface curvature of the iris, which is generally not zero. If simplifying assumptions and approximations are made about the latter factors, then a simple affine projective transformation may suffice to make the iris recognizable against itself as imaged in other poses, orthographic or not.

The essence of the problem is then estimating the two angles of gaze relative to the camera. Eye morphology is so variable in terms of visible sclera and eyelid occlusion that it is unlikely that such factors could support robust estimation, at least when only one eye is imaged; although it must be noted that humans are very impressively skilled somehow at monitoring each other’s gaze direction. In the absence of solving that mystery, an obvious alternative approach would be to assume that an orthographic image of the iris should reveal a circular pupil; therefore, detecting ellipticity of the pupil indicates off-axis image acquisition, and estimating the elongation and orientation of that ellipse would yield the two parameters of gaze deviation, modulo π in direction. We present here a somewhat more robust variant of this idea, which does not assume that the true pupil’s shape is circular when orthographically viewed. This method of estimating gaze (and, thus, correcting for off-axis imaging) uses a new approach that may be called “Fourier-based trigonometry.”

The method arises from the observation that Fourier series expansions of the X and Y coordinates of the detected pupil boundary contain shape distortion information that is related to deviated gaze, within the relationships among the real and imaginary coefficients of the lowest frequency term of each of those series expansions. In the special case that the true pupil boundary when viewed orthographically is really a circle, this method reduces to the simpler “ellipticity” method outlined above.

We begin by considering that simple special case of a circular pupil. Let $X(t)$ and $Y(t)$ be the parameterized coordinate vectors of the pupil boundary, so the range of t is from 0 to

2π in one cycle around this closed curve. Clearly, in the case of a circular pupil with radius A , which is origin centered for simplicity, these functions are just $X(t) = A \cos(t)$ and $Y(t) = A \sin(t)$. In the case of deviated gaze along a cardinal axis, and assuming that the camera distance is large compared with the iris diameter so there is simple foreshortening along the cardinal axis, these functions become $X(t) = A \cos(t)$ and $Y(t) = B \sin(t)$, where $A \neq B$. Finally, if the gaze deviation is not along a cardinal axis but rather in direction θ , then these functions take the more general conic form for an oriented ellipse, which is described as follows:

$$X(t) = [A \cos^2 \theta + B \sin^2 \theta] \cos(t) + [(B - A) \cos \theta \sin \theta] \sin(t) \quad (5)$$

$$Y(t) = [(B - A) \cos \theta \sin \theta] \cos(t) + [B \cos^2 \theta + A \sin^2 \theta] \sin(t). \quad (6)$$

It is worth noting that the information that we seek about gaze deviation, namely, the direction and magnitude of the deviation, are contained in the form of Fourier coefficients on the harmonic functions $\cos(t)$ and $\sin(t)$ that represent in their linear combination the contour data $X(t)$ and $Y(t)$. Specifically, the lowest complex-valued coefficient in a Fourier series expansion of the empirical function $X(t)$ contains as its real and imaginary parts the coefficients a and b , respectively, which specify the phase in (5) as follows:

$$a = A \cos^2 \theta + B \sin^2 \theta \quad (7)$$

$$b = (B - A) \cos \theta \sin \theta. \quad (8)$$

Likewise, the lowest complex-valued coefficient in a Fourier series expansion of the empirical function $Y(t)$ contains as its real and imaginary parts the coefficients c and d , respectively, which specify the phase in (6) as follows:

$$c = (B - A) \cos \theta \sin \theta \quad (9)$$

$$d = B \cos^2 \theta + A \sin^2 \theta. \quad (10)$$

Thus, we can derive the gaze deviation parameters that we seek just by computing the relevant Fourier coefficients of the empirical contour functions $X(t)$ and $Y(t)$. This estimation process is independent of the higher order Fourier coefficients that will exist when the pupil has a shape that is more complex and irregular than a circle. The method is not restricted to such an assumption about a circular shape.

Algebraic manipulation extracts from the four empirical Fourier coefficients a , b , c , and d the gaze deviation parameters that we need. It should be noted that although the right-hand sides of (8) and (9) above appear to be identical functions of the desired parameters, these equations express constraints based on different empirical data. Quantities a and b are obtained from $X(t)$, whereas c and d are obtained from $Y(t)$. The computed direction of gaze deviation (modulo π) essentially has the form of Fourier phase information, i.e.,

$$\theta = 0.5 \arctan \left(\frac{-b - c}{a - d} \right) \quad (11)$$

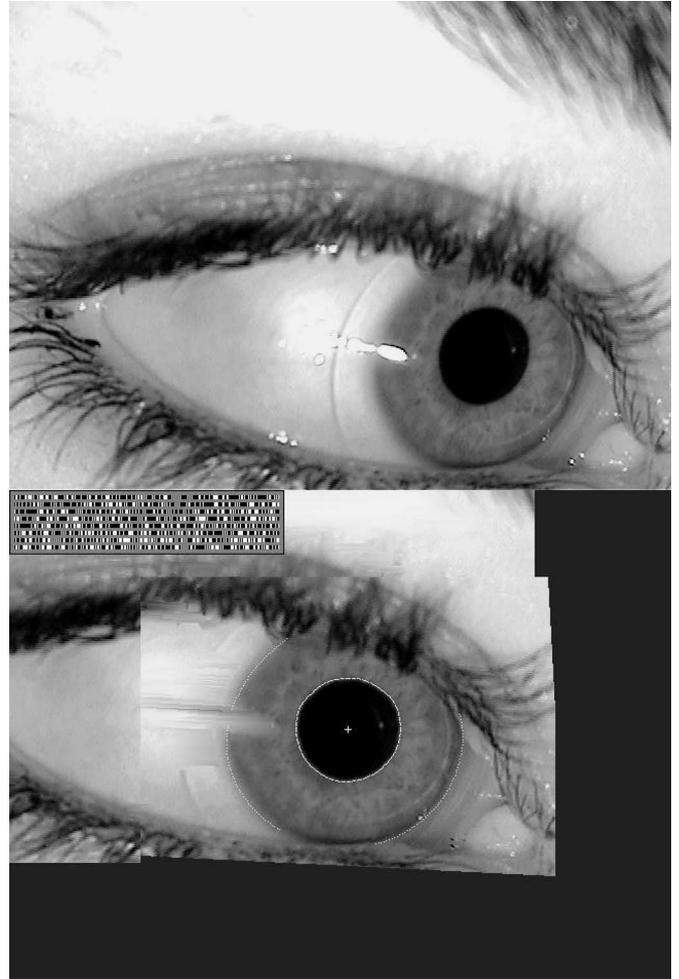


Fig. 3. Fourier-based trigonometry enables gaze estimation and transformation of an image of an eye with deviated gaze into one apparently looking directly at the camera. Without this transformation, such images would fail to be matched. This iris is ICE-1 file 244858.

and the magnitude of the gaze deviation in direction θ is expressed not as an angle but as the projective aspect ratio $\gamma = B/A$, which gives the second affine transformation parameter as

$$\gamma = \frac{(a + d) \cos(2\theta) + a - d}{(a + d) \cos(2\theta) - a + d}. \quad (12)$$

By estimating these parameters, the “Fourier-based trigonometry” allows the projective geometric deformation caused by the gaze deviation to be reversed by an affine transformation of the off-axis image. This is illustrated in Fig. 3, which shows image 244858 from the NIST ICE-1 database in the upper panel and the same eye image after “correcting” the gaze deviation by an affine transformation with the extracted parameters (θ, γ) in the lower panel. The result of the transformation is to convert such images into apparent orthographic form, appearing to rotate the eyes in their sockets, making them recognizable against other images of the same eye. A limitation of this method is that the affine transformation assumes that the iris is planar, whereas, in fact, it has some curvature.

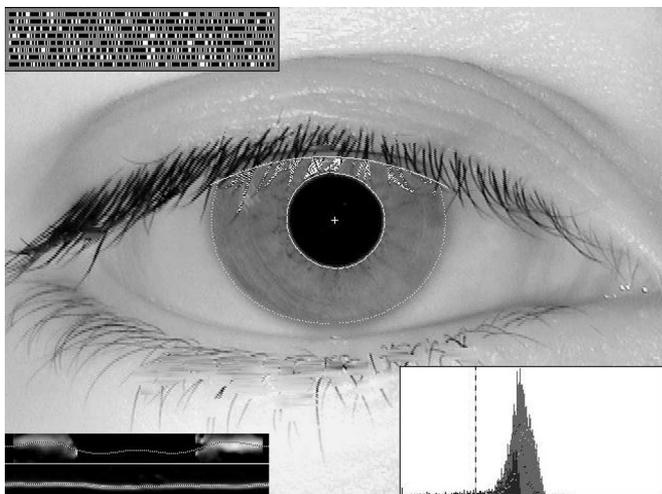


Fig. 4. Statistical inference of eyelashes from the iris pixel histogram and determination of a threshold for excluding the lashes (labeled in white) from influencing the IrisCode. This iris is ICE-1 file 239766.

IV. EXCLUDING EYELASHES BY STATISTICAL INFERENCE

One of the ways in which iris image data may be corrupted, besides reflections, camera noise, and eyelid occlusion, is occlusion by eyelashes (usually from the upper eyelid). These often have random and complex shapes, combining with each other to form masses of intersecting elements rather than just simple hair-like strands that might be amenable to detection by elementary shape models. They can be the strongest signals in the iris image in terms of contrast or energy, and they could dominate the IrisCode with spurious information if not detected and excluded from the encoded data.

The inference of eyelashes and their exclusion from the IrisCode can be handled by statistical estimation methods that essentially depend on determining whether the distribution of iris pixels is multimodal. If the lower tail of the iris pixel histogram supports a hypothesis of multimodal mixing, then an appropriate threshold can be computed, and pixels outside it can be excluded from influencing the IrisCode.

Fig. 4 illustrates this method. The panel in its lower-right corner superimposes four histograms, all computed from just the pixels in the segmented iris portion of the image between the detected eyelids. The solid gray distribution is a histogram of all the iris pixels (ranging from 0 to 255). The two dotted outline histograms break this up into two components, one for just the lower part of the iris (white dotted curve), and the other for just the upper part of the iris (black dotted curve). The solid black histogram is the difference between these two histograms. We are interested in whether the cumulatives from the left of this difference histogram pass a test of being statistically separable from the mother (solid gray) distribution. If so, based on significant Z -score deviations between their respective quartiles, the hypothesis that the iris contains some superimposed eyelash pixels may be accepted.

The vertical dashed line in the histogram panel indicates the computed threshold where such a hypothesis in this case (for this image) is supported. If that hypothesis also passes a further test on the deviation between the threshold quartile and

the median of the mother distribution, thus confirming that not only are there two populations but also that they are sufficiently different from each, then the threshold is accepted, and pixels below it are deemed to arise from superimposed eyelashes.

In the iris image itself in Fig. 4, these detected eyelashes within the iris have now been marked as white pixels. Their positions are recorded in a mask array that prevents them from influencing the data that encode the iris texture.

V. SCORE NORMALIZATIONS

Iris recognition works by performing a test of statistical independence between two IrisCodes, in order to decide whether they arise from the same or from different irises [2]–[4]. This test of statistical independence is equivalent to tossing a coin many times, with each toss representing a comparison between two bits in the two IrisCodes, in order to decide whether or not the coin is fair by delivering roughly 50–50 outcomes. If such a result is obtained, then the irises can be judged as independent; but if there is a great preponderance of, e.g., “heads,” which means that a large majority of corresponding bit pairs agreed, then that is strong evidence that the IrisCodes came from the same iris. But what is the effect of greatly varying numbers of such “coin tosses” in these correlated Bernoulli trials, due to varying amounts of iris data being visible and available for comparison?

Areas of the iris that are obscured by eyelids, eyelashes, or reflections from eyeglasses, or that have low contrast or a low signal-to-noise ratio, are detected by the image-processing algorithms and prevented from influencing the iris comparisons through bitwise mask functions. Whereas the IrisCode bits themselves contain phase data [2]–[4] that are XORed (\otimes) to detect disagreement and thereby determine similarity between two irises, the bits to be considered are first selected by ANDing (\cap) each pair with the associated mask functions of both irises to ensure their validity and their significance. The norms ($\| \cdot \|$) of the resultant bit vector and of the ANDed mask vectors are then measured in order to compute a raw Hamming distance HD_{raw} , as the fraction of meaningful bits that disagree between two irises whose two phase code bit vectors are denoted $\{\text{codeA}, \text{codeB}\}$ and whose mask bit vectors are denoted $\{\text{maskA}, \text{maskB}\}$. Thus, we have

$$HD_{\text{raw}} = \frac{\|(\text{codeA} \otimes \text{codeB}) \cap \text{maskA} \cap \text{maskB}\|}{\|\text{maskA} \cap \text{maskB}\|}. \quad (13)$$

Usually, the total number of bit pairs that are available for comparison, i.e., $\|\text{maskA} \cap \text{maskB}\|$, is nearly a thousand. However, if one of the irises has, for example, almost complete occlusion of its upper half by a drooping upper eyelid, and if the other iris that is being compared with it has almost complete occlusion of its lower half, then the common area available for comparison may be almost nil. In such cases, returning to the coin-tossing analogy, our test for the “fairness” of the coin (i.e., statistical independence of the two IrisCodes by finding a nearly 50–50 result) will be based upon a very small number of Bernoulli trials indeed. Therefore, the interpretation of any given deviation from the 50–50 outcome that is expected

for independence must take into account the total amount of comparison data that were available. This is the role of score normalization.

A natural choice for the score normalization rule would be to rescale all deviations from a 0.5 raw Hamming distance in proportion to the square root of the number of bits that were compared when obtaining that score. The reason for such a rule is that the expected standard deviation in the distribution of Bernoulli trial outcomes (expressed as a fraction of the n Bernoulli trials having a given outcome) is $\sigma = (pq/n)^{1/2}$, where p and q are the respective outcome probabilities (both nominally 0.5 in this case). Thus, decision confidence levels can be held constant, irrespective of how many bits n were actually compared, by mapping each raw Hamming distance HD_{raw} into a normalized score HD_{norm} using a rescaling rule like

$$HD_{\text{norm}} = 0.5 - (0.5 - HD_{\text{raw}}) \sqrt{\frac{n}{911}}. \quad (14)$$

This normalization should transform all samples of scores obtained when comparing different eyes into samples drawn from the same binomial distribution, whereas the raw scores HD_{raw} might be samples from many different binomial distributions having standard deviations σ that are dependent on the number of bits n that were actually available for comparison. This normalization maintains constant confidence levels for decisions using a given Hamming distance threshold, regardless of the value of n . The scaling parameter 911 is the typical number of bits compared (unmasked) between two different irises, as estimated from one particular (early) database. Today, based on much larger databases, it appears that a better choice for that parameter may be about 960. In any case, when comparing different normalization rules in this paper, this rule (14) will be termed “SQRT normalization.”

The benefit of SQRT normalization is to prevent false matches arising by chance due to few bits being compared (just as few coin tosses may well yield all “heads”). However, the cost of this normalization for n is that same-eye matches are penalized when few bits are available for comparison; even if they all agreed, so that $HD_{\text{raw}} = 0$, the resulting HD_{norm} may be above the acceptance threshold, and the match would be rejected. This penalty is apparent by comparing false reject performance with and without score normalization on the NIST ICE-1 iris database, which consists of a few thousand iris images that NIST released together with “ground-truth” information. This image database contained many very difficult and corrupted images, often in poor focus and with much eyelid occlusion, and sometimes with the iris partly outside the image frame. In the region of the receiver operating characteristic (ROC) curve [tradeoff between FRR and false accept rate (FAR)], where one tolerates rather high FAR such as one in 1000 or one in 10 000, as shown in Fig. 5, the cost of score normalization (Algorithm 1) on FRR is clear. The EER (where $FRR = FAR$) is about 0.001 without score normalization, but 0.002 with normalization. Similarly, at other nominal points of interest in this region of the ROC curve, as tabulated within Fig. 5, the cost of score normalization is roughly a doubling in FRR.

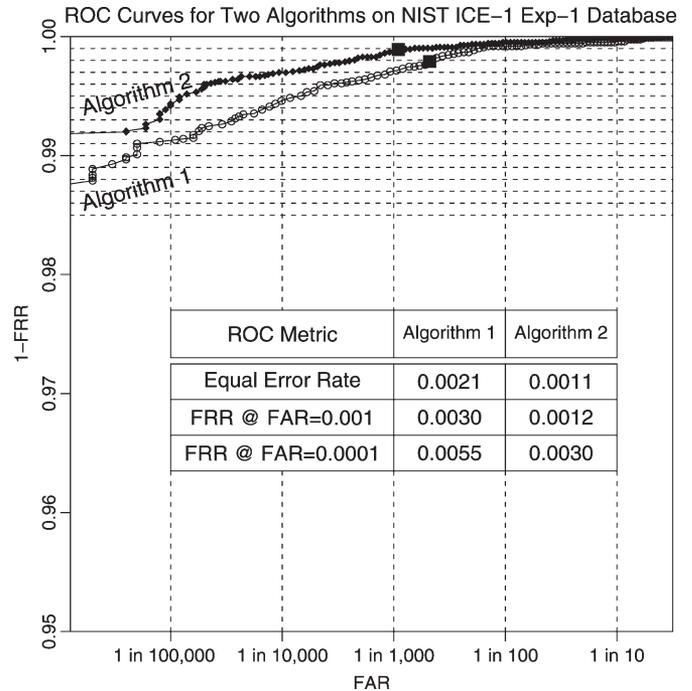


Fig. 5. Comparison of two algorithms on the NIST ICE-1 database: with and without score normalization (i.e., Algorithms 1 and 2, respectively). Solid squares mark their EER points. In this region of the ROC curve near EERs, where the required FMRs are not very demanding, best performance is achieved without score normalization.

However, in much more aggressive regions of the ROC, where one demands an FAR of perhaps one in a billion for applications such as national database search or “all-against-all” cross-comparisons over large databases to discover any multiple IDs, the benefits of score normalization become apparent. It is informative to first examine what occurs in this region *without* score normalization. We must use a much larger database than the released part of the NIST ICE-1 database, because the latter allows only a few million cross-comparisons between different eyes. We must explore the behavior when many billions of cross-comparisons are done. This is possible with the United Arab Emirates (UAE) database of $N = 632\,500$ IrisCodes all obtained from different eyes, enabling a total of $N \cdot (N - 1)/2 = 200$ billion different pair comparisons. Based on this complete set of 2×10^{11} pairings, Table I shows the observed FMRs for unnormalized scores HD_{raw} as a function of both the decision threshold and the number of bits compared.

It is clear from Table I that false matches are much more likely (at any Hamming distance criterion) when fewer bits were the basis of the comparison. For example, at a criterion of $HD_{\text{raw}} = 0.285$, using unnormalized scores, a false match would be 1000 times more likely when only 400 bits were the basis of the comparison than if 1000 bits were available for comparison. This illustrates the importance of score normalization.

For the same UAE database, Fig. 6 compares the FMRs for the three approaches to normalizing scores based on the number of bits compared. The upper plot takes no account of the number of bits, so it has the highest FMRs. (It should be

TABLE I
FMR WITHOUT SCORE NORMALIZATION: DEPENDENCE ON THE NUMBER OF BITS COMPARED AND THE CRITERION

HD _{Crit}	400 bits	500 bits	600 bits	700 bits	800 bits	900 bits	1000 bits
0.260	$2 \cdot 10^{-9}$	$5 \cdot 10^{-10}$	$3 \cdot 10^{-10}$	$1 \cdot 10^{-10}$	0	0	0
0.265	$3 \cdot 10^{-9}$	$8 \cdot 10^{-10}$	$5 \cdot 10^{-10}$	$2 \cdot 10^{-10}$	$4 \cdot 10^{-11}$	0	0
0.270	$4 \cdot 10^{-9}$	$1 \cdot 10^{-9}$	$9 \cdot 10^{-10}$	$5 \cdot 10^{-10}$	$2 \cdot 10^{-10}$	0	0
0.275	$7 \cdot 10^{-9}$	$2 \cdot 10^{-9}$	$1 \cdot 10^{-9}$	$9 \cdot 10^{-10}$	$5 \cdot 10^{-10}$	$3 \cdot 10^{-11}$	0
0.280	$1 \cdot 10^{-8}$	$4 \cdot 10^{-9}$	$2 \cdot 10^{-9}$	$2 \cdot 10^{-9}$	$1 \cdot 10^{-9}$	$2 \cdot 10^{-10}$	0
0.285	$2 \cdot 10^{-8}$	$7 \cdot 10^{-9}$	$4 \cdot 10^{-9}$	$3 \cdot 10^{-9}$	$2 \cdot 10^{-9}$	$5 \cdot 10^{-10}$	$2 \cdot 10^{-11}$
0.290	$3 \cdot 10^{-8}$	$1 \cdot 10^{-8}$	$8 \cdot 10^{-9}$	$7 \cdot 10^{-9}$	$4 \cdot 10^{-9}$	$1 \cdot 10^{-9}$	$1 \cdot 10^{-10}$
0.295	$4 \cdot 10^{-8}$	$2 \cdot 10^{-8}$	$1 \cdot 10^{-8}$	$1 \cdot 10^{-8}$	$9 \cdot 10^{-9}$	$3 \cdot 10^{-9}$	$4 \cdot 10^{-10}$
0.300	$6 \cdot 10^{-8}$	$3 \cdot 10^{-8}$	$3 \cdot 10^{-8}$	$2 \cdot 10^{-8}$	$2 \cdot 10^{-8}$	$7 \cdot 10^{-9}$	$9 \cdot 10^{-10}$
0.305	$9 \cdot 10^{-8}$	$6 \cdot 10^{-8}$	$5 \cdot 10^{-8}$	$4 \cdot 10^{-8}$	$4 \cdot 10^{-8}$	$1 \cdot 10^{-8}$	$2 \cdot 10^{-9}$
0.310	$1 \cdot 10^{-7}$	$1 \cdot 10^{-7}$	$8 \cdot 10^{-8}$	$8 \cdot 10^{-8}$	$7 \cdot 10^{-8}$	$3 \cdot 10^{-8}$	$5 \cdot 10^{-9}$
0.315	$2 \cdot 10^{-7}$	$2 \cdot 10^{-7}$	$1 \cdot 10^{-7}$	$2 \cdot 10^{-7}$	$1 \cdot 10^{-7}$	$6 \cdot 10^{-8}$	$1 \cdot 10^{-8}$
0.320	$3 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$2 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$1 \cdot 10^{-7}$	$2 \cdot 10^{-8}$

False Match Rate vs Criterion (200 Billion Cross-comparisons)

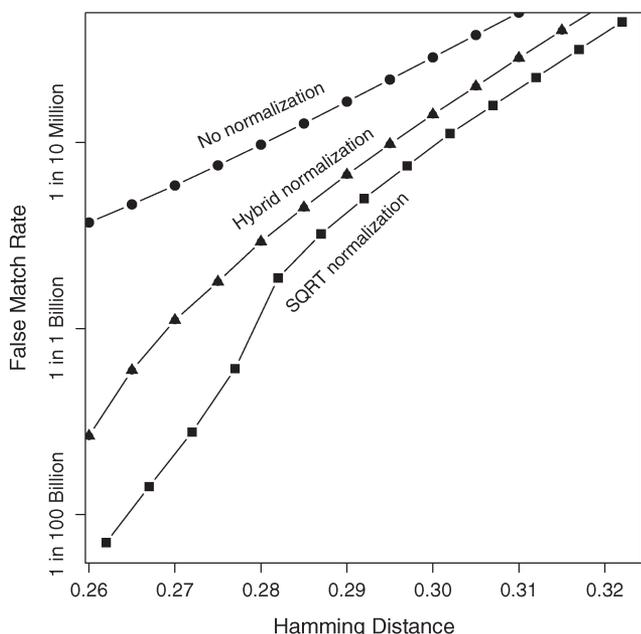


Fig. 6. Comparisons of three approaches to score normalization based on 200 billion iris cross-comparisons using the UAE database of 632,500 IrisCodes. FMRs are plotted in semilogarithmic coordinates versus decision criterion. The range of the ordinate in this plot spans a factor of 300,000 to 1.

noted that, in Fig. 6, all iris comparisons were done in each of several relative orientations to compensate for possible tilt, making all FMRs about seven times worse than the case in Table I because of the increased number of opportunities.) The bottom curve in Fig. 6 shows the benefit of score normalization using the SQRT rule given in (14), causing the observed FMR to plummet to one in 200 billion at a criterion around $HD_{norm} = 0.262$. This performance is about 2000 times better than without the normalization (note the semilogarithmic coordinates). The middle curve represents a hybrid normalization rule that is a linear combination of the other two, taking into account the number of bits compared only when in a certain range.

We learn from these comparisons of alternative normalization rules that performance in different regions of the ROC curve is optimized by different rules. In the relatively undemanding domain highlighted in Fig. 5, near the EER point

Score Distribution for 200 Billion Different Iris Comparisons

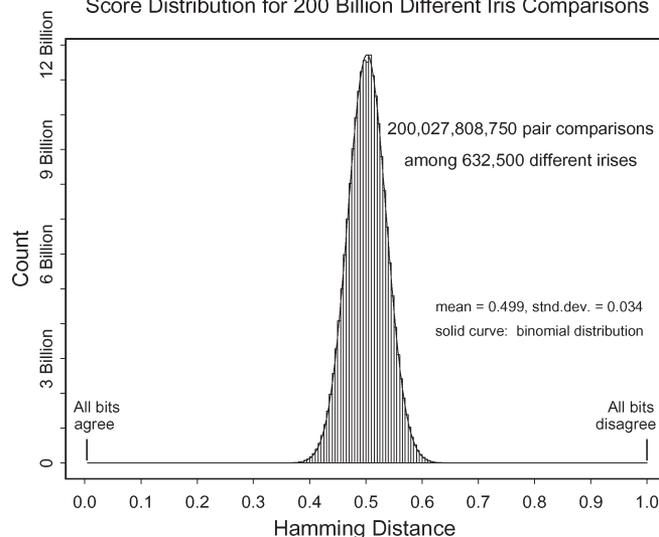


Fig. 7. Distribution of HD_{norm} normalized similarity scores (14) for 200 billion different pairings of iris patterns, without relative rotations. The solid curve fitting the histogram is a binomial density (15).

of the ROC curve and involving only a small database, best performance was achieved without score normalization because that amounts to a penalty on good matches when few bits were compared. However, in the vastly more demanding domain of national-scale databases that involve possibly astronomical numbers of cross-comparisons, as investigated in Fig. 6 with 200 billion iris comparisons, normalizing scores by the number of bits on which they were based is a critical necessity.

VI. ADAPTING FOR LARGE-SCALE DATABASES

Using the SQRT normalization rule, Fig. 7 presents a histogram of all 200 billion cross-comparison similarity scores among the 632,500 different irises in the UAE database. The vast majority of IrisCodes from different eyes disagreed in roughly 50% of their bits, as expected, since the bits are equiprobable and uncorrelated between different eyes [2], [4]. Very few pairings of IrisCodes could disagree in fewer than 35% or more than 65% of their bits, as is evident from the distribution.

The solid curve that fits the data very closely in Fig. 7 is a binomial probability density function. This theoretical form was chosen because comparisons between bits from different IrisCodes are Bernoulli trials, or conceptually “coin tosses,” and Bernoulli trials generate binomial distributions. If one tossed a coin whose probability of “heads” is p in a series of N independent tosses and counted the number m of “heads” outcomes, and if one tallied this fraction $x = m/N$ in a large number of such repeated runs of N tosses, then the expected distribution of x would be as per the solid curve in Fig. 7. Thus, we have

$$f(x) = \frac{N!}{m!(N-m)!} p^m (1-p)^{(N-m)}. \quad (15)$$

The analogy between tossing coins and comparing bits between different IrisCodes is deep but imperfect, because any given IrisCode has internal correlations arising from iris features, especially in the radial direction [2]. Further correlations are introduced because the patterns are encoded using 2-D Gabor wavelet filters [4], whose low-pass aspect introduces correlations in amplitude, and whose bandpass aspect introduces correlations in phase, both of which linger to an extent that is inversely proportional to the filter bandwidth. The effect of these correlations is to reduce the value of the distribution parameter N to a number that is significantly smaller than the number of bits that are actually compared between two IrisCodes; N becomes the number of effectively independent bit comparisons. The value of p is very close to 0.5 (empirically 0.499 for the UAE database), because the states of each bit are equiprobable *a priori*, and so any pair of bits from different IrisCodes are equally likely to agree or disagree.

The binomial functional form that very well describes the distribution of normalized similarity scores for comparisons between different iris patterns is the key to the robustness of these algorithms in large-scale search applications. The tails of the binomial attenuate extremely rapidly, because of the dominating central tendency caused by the factorial terms in (15). Rapidly attenuating tails are critical for a biometric to survive the vast numbers of opportunities to make false matches without actually making any, when applied in an “all-against-all” mode of searching for matching or multiple IDs, as contemplated in some national ID card projects in the U.K. and elsewhere in Europe.

The cumulatives (up to various thresholds) under the left tail of the distribution of normalized similarity scores for different irises compared at multiple relative tilts reveal the FMRs among the 200 billion iris comparisons if the identification decision policy used those thresholds. These rates are provided in Table II. Although the smallest observed match was around 0.26, the table has been extended down to 0.22 using the theoretical cumulative of the extreme value distribution of multiple samples from the binomial (15) plotted as the solid curve in Fig. 5 in order to extrapolate the theoretically expected FMRs for such decision policies. These FMRs, whether observed or theoretical, also serve as confidence levels that can be associated with a given quality of match using the score normalization rule (14). In this analysis, only a single eye is presumed to be

TABLE II
FMRs WITH HD_{norm} SCORE NORMALIZATION: DEPENDENCE ON THE CRITERION (200 BILLION COMPARISONS, UAE DATABASE)

<i>HD Criterion</i>	<i>Observed False Match Rate</i>
0.220	0 (<i>theor: 1 in 5×10^{15}</i>)
0.225	0 (<i>theor: 1 in 1×10^{15}</i>)
0.230	0 (<i>theor: 1 in 3×10^{14}</i>)
0.235	0 (<i>theor: 1 in 9×10^{13}</i>)
0.240	0 (<i>theor: 1 in 3×10^{13}</i>)
0.245	0 (<i>theor: 1 in 8×10^{12}</i>)
0.250	0 (<i>theor: 1 in 2×10^{12}</i>)
0.255	0 (<i>theor: 1 in 7×10^{11}</i>)
0.262	1 in 200 billion
0.267	1 in 50 billion
0.272	1 in 13 billion
0.277	1 in 2.7 billion
0.282	1 in 284 million
0.287	1 in 96 million
0.292	1 in 40 million
0.297	1 in 18 million
0.302	1 in 8 million
0.307	1 in 4 million
0.312	1 in 2 million
0.317	1 in 1 million

presented. Under the assumption of independence between the right- and left-eye IrisCodes, which is strongly supported by the available data (see [4, Fig. 6]), the confidence levels in Table II could be multiplied together for matches that were obtained with both eyes.

The requirements of biometric operation in the “identification” mode by exhaustively searching a large database are vastly more demanding than merely operating in a one-to-one “verification” mode (in which an ID must first be explicitly asserted, which is then verified in a yes/no decision by comparison against just the single nominated template).

If P_1 is the false match probability for single one-to-one verification trials, then $(1 - P_1)$ is the probability of not making a false match in single comparisons. The likelihood of successfully avoiding this in each of N independent attempts is therefore $(1 - P_1)^N$; hence, P_N , which is the probability of making at least one false match when searching a database containing N different patterns, is given by

$$P_N = 1 - (1 - P_1)^N. \quad (16)$$

By observing the approximation that $P_N \approx NP_1$ for small $P_1 \ll 1/N \ll 1$, when searching a database of size N , an identifier needs to be roughly N times better than a verifier to achieve comparable odds against making false matches. In effect, as the database grows larger and larger, the chance of making a false match also grows almost in proportion. These chances also grow in proportion to the number of independent searches that are conducted against the database. To successfully survive so many opportunities to make false matches, the decision threshold policy must be adaptive to both of these factors. Fortunately, because of the underlying binomial combinatorics, the algorithms with score normalization generate extremely rapidly attenuating tails for the HD_{norm} distribution. This felicitous property enables very large databases to be

accommodated and large numbers of searches to be conducted against them. The rule to be followed for decision policy threshold selection is to multiply the size of the enrolled database times the number of searches to be conducted against it in a given interval of time and then to determine from Table II what Hamming distance threshold will correspond to the risk level that is deemed to be acceptable. Again, Table II assumes that only a single eye is used for matching.

For example, in the U.K., with a national adult population of about 45 million, an “all-against-all” comparison of IrisCodes, which totals to about 10^{15} pairings, as envisioned to search for any multiple IDs when issuing the proposed biometric ID cards, could be performed using an HD_{norm} decision threshold as high as 0.22 without expecting to make any false matches. At this threshold, in order to keep the FnMR acceptably low (e.g., below 1%), systems will need to be more tolerant than the currently deployed ones in order to handle difficult image capture conditions, or unusual eyes and nonideal presentations. It is hoped that the new image-processing algorithms introduced in Sections II–IV of this paper using active contours, deviated gaze correction, and other improvements can help in achieving those requirements. A further advance that allows IrisCodes to be *indexed* by their collisions with substrings, thereby replacing the need for exhaustive search (which is time consuming in large databases) by instead almost instantaneous direct addressing with IrisCodes, is the subject of a separate paper [6].

REFERENCES

- [1] A. Blake and M. Isard, *Active Contours*. Heidelberg, Germany: Springer-Verlag, 1998.
- [2] 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.
- [3] J. G. Daugman, “Biometric personal identification system based on iris analysis,” U.S. Patent 5,291,560, Mar. 1, 1994. U.S. Pat. Off., Washington, DC.

- [4] J. G. Daugman, “How iris recognition works,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 21–30, Jan. 2004.
- [5] V. Dorairaj, N. Schmid, and G. Fahmy, “Performance evaluation of non-ideal iris based recognition system implementing global ICA encoding,” in *Proc. IEEE ICIP*, 2005, vol. 3, pp. 285–288.
- [6] F. Hao, J. G. Daugman, and P. Zielinski, “A fast search algorithm for a large fuzzy database,” *IEEE Trans. Inf. Forensics Security*, 2007.
- [7] *Information Technology—Biometric Data Interchange Formats, Part 6: Iris Image Data*, 2005. ISO/IEC Standard 19794-6.
- [8] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.
- [9] J. Matey, K. Hanna, R. Kolczynski, D. LoIacono, S. Mangru, O. Naroditsky, M. Tinker, T. Zappia, and W-Y. Zhao, “Iris on the move: Acquisition of images for iris recognition in less constrained Environments,” *Proc. IEEE*, vol. 94, no. 11, pp. 1936–1947, Nov. 2006.
- [10] Website about generalized Faberge embeddings. [Online]. Available: <http://users.vnet.net/schulman/Faberge/eggs.html>



John Daugman received the A.B. and Ph.D. degrees from Harvard University, Cambridge, MA.

He was with the faculty of Harvard University before he joined Cambridge University, Cambridge, U.K., in 1991, where he is currently with the Computer Laboratory. He has held visiting Professorships at the University of Groningen, Groningen, The Netherlands, as the Johann Bernoulli Chair of Mathematics and Informatics, and at the Tokyo Institute of Technology, Tokyo, Japan, as the Toshiba Endowed Chair. His areas of research and teaching are computer vision and biological vision, pattern recognition, information theory, and neural computing. He is the inventor of iris recognition, and his algorithms currently underlie all public deployments of this technology worldwide. He serves or has served as an Associate Editor of *Network: Computation in Neural Systems*, the *International Journal of Wavelets, Multiresolution, and Information Processing*, *Cognitive Brain Research*, and the *Journal of Computation and Mathematics*.

Dr. Daugman has served as an Associate Editor of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. He received the U.S. Presidential Young Investigator Award from the National Science Foundation, the Information Technology Award and Medal from the British Computer Society, the “Millennium Product” Award from the U.K. Design Council, the Smithsonian and “Time 100” Innovators Award (U.S.), and the OBE, Order of the British Empire, for his scientific and technical work.